NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

THESIS

ANALYSIS OF AN IMPERFECT INFORMATION FLOW REDUCTION AND SORTING SYSTEM

by

John A. Sears, III

March, 1997

Thesis Advisor: Kneale T. Marshall

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE March, 1997	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE ANALYSIS OF AN IMPERFECT INFORMATION FLOW REDUCTION AND SORTING SYSTEM			5. FUNDING NUMBERS
6. AUTHOR(S) SEARS, John A., III			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey CA 93943-5000			8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.			12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

This thesis studies the employment of an information flow reduction and sorting system. The system is designed to reduce the amount of information gathered by a collection system to a rate that a user of that information can accept. The thesis demonstrates the benefits of trait-based analysis of information as a method of screening desired information from undesired. These systems increase the quality of the information reaching the user while adding a delay to achieve the screening process. A method of networking the screening devices is discussed. A sorting system is added to the screening process to demonstrate its ability to increase the speed of desired information through the system. The models are illustrated through numerical examples. The analysis provides the user of these systems with an understanding of their design, employment, benefits, costs and calibration requirements.

14. SUBJECT TERMS INFORMATION SYSTEM, INFORMATION REDUCTION, INFORMATION SORTING, QUEUEING			15. NUMBER OF PAGES 94
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18 298-102

Approved for public release; distribution is unlimited.

ANALYSIS OF AN IMPERFECT INFORMATION FLOW REDUCTION AND SORTING SYSTEM

John A. Sears, III
Lieutenant Commander, United States Navy
B.S., Pennsylvania State University, 1985

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1997

ABSTRACT

This thesis studies the employment of an information flow reduction and sorting system. The system is designed to reduce the amount of information gathered by a collection system to a rate that a user of that information can accept. The thesis demonstrates the benefits of trait-based analysis of information as a method of screening desired information from undesired. These systems increase the quality of the information reaching the user while adding a delay to achieve the screening process. A method of networking the screening devices is discussed. A sorting system is added to the screening process to demonstrate its ability to increase the speed of desired information through the system. The models are illustrated through numerical examples. The analysis provides the user of these systems with an understanding of their design, employment, benefits, costs and calibration requirements.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Information collection systems gather information from the environment and deliver it to a decision maker. The decision maker's goal is to evaluate this information and then act upon it based upon what he/she has learned. The decision maker has conflicting requirements for the collection system: high quality information and high volume of information. These requirements conflict because the decision maker's enemy desires to counter the collection efforts. The enemy obscures vital information from the decision maker by mixing it with non-vital information. This decreases the quality of the collected information. The enemy also hides vital information through various means, allowing the collection system to obtain only a small portion of what the decision maker desires. To obtain the desired amount of vital information, the decision maker is forced to collect an increasingly larger amount of raw information. Often this level exceeds the decision maker's ability to analyze all of the collected items. The role of the decision maker, therefore, is to collect as much of the available information as possible, remove the non-vital information and make the correct decision based upon the obtained vital information in an acceptable amount of time. It is to the decision maker's advantage to possess an automated system which reduces a high volume of collected information to a useable level while increasing the overall quality.

This thesis studies the use of a trait-based information reduction system designed to achieve the flow reduction and quality improvement desired by the decision maker. The reduction system examines traits of an arriving piece of information to determine if those traits are similar to those of vital information. The system forms acceptance/rejection criteria based upon these traits and forwards information to the decision maker at a rate the

decision maker can accept. An example demonstrates quality level improvements of 150 percent for a single trait system and 170 percent for a two trait system. These improvements are not without cost. The costs associated with trait-based selection systems include infrastructure costs such as design and fielding as well as time delay costs associated with the trait analysis. The example demonstrates a 10 percent increase in analysis time for a single trait system and a 14 percent increase for a two trait system. A method of calibrating trait-based reduction systems to achieve optimal quality is provided. Sensitivity of these systems to changes in the information environment is demonstrated.

It is desirable to reduce the time delay associated with a trait-based reduction system so that the decision maker receives the vital information sooner. This thesis also studies the benefits of adding a trait-based sorter to a network of trait-based selectors. The sorter is calibrated to give vital information priority over non-vital information. The thesis finds that a sort system of this type always provides a decrease in time delay of vital information provided to the decision maker. The example demonstrates that a sorter added to the single trait reduction system decreased the overall delay by 15 percent. A method for obtaining the optimal decrease in delay is provided as well as a discussion of the sensitivity of the system to changes in the information environment.

I. INTRODUCTION

Information is the key ingredient in a decision making process. The structure of military organizations has a basis in the flow of information from those who collect it to those who determine the response to it. As the sophistication and capacity of information collection technology continues to grow, the amount of information presented to the decision maker increases. At face value this is a desired effect. However, the vast assortment of available facts presented to today's decision makers can easily overwhelm his/her capability to absorb the material. Equally sophisticated information reduction and sorting systems attempt to remove the chaff from the wheat and present the decision maker with the most relevant facts.

Reduction and sorting systems that deal with such issues are currently in use throughout the military. An example might be as warfighting oriented as a Combat Information Center, where sensor operators utilize their collection systems to ascertain the battle environment. The task of these operators is to present the best picture to the Tactical Action Officer (TAO) for display as well as to provide assessments on particular portions of that information. The TAO must then make tactical decisions based upon the presented information. National level agencies collect information in a similar fashion and with the same underlying purpose: to provide a decision maker with as many available and relevant facts and assessments on the issue under evaluation in the available time.

Systems of this type possess inherent problems that confound the decision process. Received information is likely to be incomplete or shrouded from its absolute meaning. The cause may be due to limitations in the collection system. The radar holds an air contact but is it military or commercial? Is it friend or foe? Active denial of information as well as infusion of false information by the target increases the chance that the

assessment will be incorrect. This particularly applies to military and national level information. Systems and procedures designed to respond to clarify the vague information often require the application of assets and time. The TAO may require the intercept of the air contact for visual identification. The intelligence officer may require the decryption of the enciphered information. This delay may allow the target to operate within the decision maker's decision cycle and afford the enemy tactical or strategic advantage. Incorrect assessment of the information also negatively impacts the decision by providing a falsely high or falsely low evaluation of the threat. Allowing either to occur is detrimental to the decision maker's ability to effectively respond.

The objective of this thesis is to provide a model for the analysis of an imperfect information flow reduction and sorting system. Envision the system as a network of sorting and analysis devices, Figure 1, within which items of information are placed into priority levels according to discrete features of the items. The resultant stack is of items in prioritized order presented to a decision maker. The decision maker does not have the capability to review the entire stack but will review the objects from highest to lowest level until he must make a decision. The system is imperfect due to its inability to always place an object within the proper priority level. Three results with two types of errors can exist during the sorting of each object.

1. The object receives the correct priority level. No error occurred.

2. The object receives a lower priority than it actually holds. This reduces the quality of the objects available to the decision maker or delays the time of presentation.

3. The object receives a higher priority than it actually holds. This confounds the decision process by placing some chaff with the wheat. It also lowers the efficiency of the system.

The problem is further defined in Chapter II. The models are developed in Chapter III using a simple reduction system as a base-case. Section E of Chapter III discusses the addition of a single measurement and

selector pair. Section F links an additional pair and notes the implication of networking such pairs. Section G develops the sort system model as an enhancement to a single selector model. Chapter IV discusses the calibration of these devices and their sensitivity to changes in the inputs. Chapter V provides insight on the models and recommendations for further analysis.

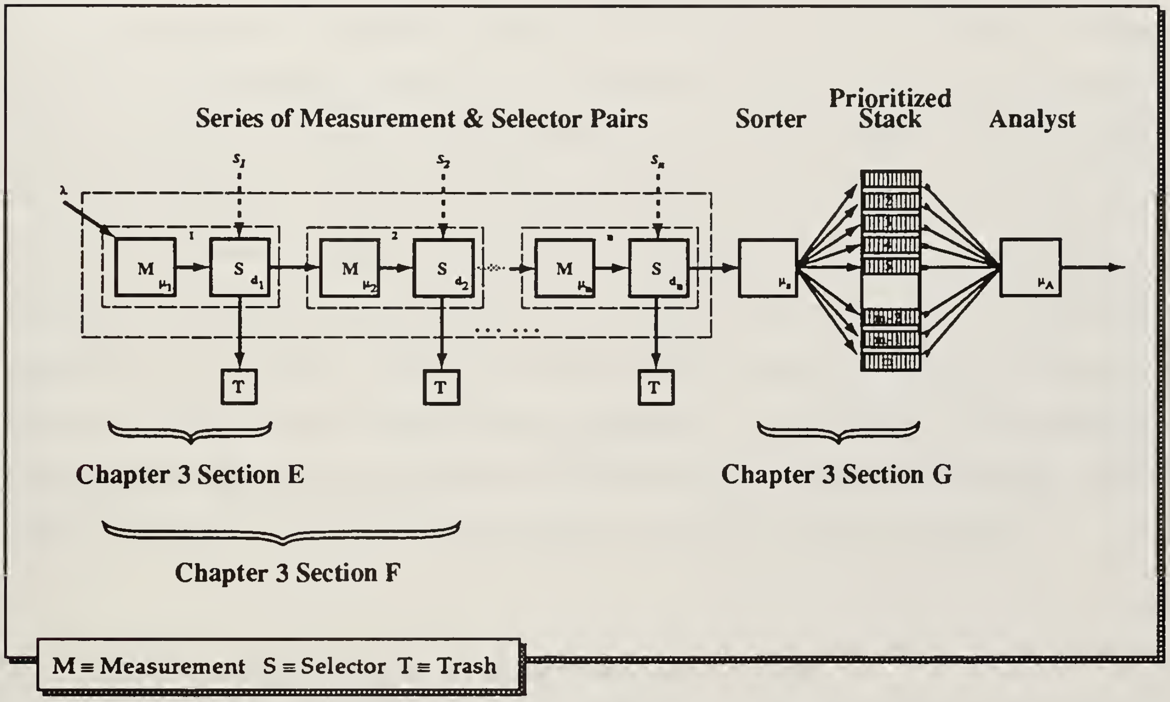

Figure 1. Generic System Illustration with Model Development Notation

II. PROBLEM DISCUSSION

A. GENERAL SYSTEM MODELS

Information collection systems gather information of various forms from the environment and pass it to a decision maker. The decision maker wishes to make the best decision possible with the available information in the available time. Often, the collection systems can collect more information than the decision maker can review and process. This causes a backlog in available information and results in a decision maker who is constantly analyzing information and not using his/her time to make the required decisions. The result is information overload. Information flow reduction systems seek to lessen the amount of information reaching a decision maker while, if possible, increasing the quality of the information available.

1. Simple Information Flow Reduction Systems

The simplest method of flow reduction is to require the amount of information passed to the decision maker to be less than a previously set threshold. The decision maker must set this limit based upon his/her own capabilities. Simple information flow reduction systems make discriminating decisions on which items to reject and which to retain. Three methods exist for achieving this type of flow reduction:

- rejecting previously collected information,
- reducing the amount of information gathered by the collection system and,
- reshaping the environment from which information is obtained.

This thesis analyzes the flow reduction that results from the use of a simple rejection procedure. The thesis does not examine the reduction in the amount of information gathered by the collection system. Section 4 contains a discussion regarding the reshaping of the information environment.

2. More Complex Information Flow Reduction Systems

A more complex method of reducing the flow of information to a decision maker is to allow for devices between the collection system and the decision maker to pre-analyze the data. These pre-analyzers seek to remove the same level of flow as the simple reduction system. However, the pre-analyzers have the additional capability of examining traits of the information and determining acceptance and rejection based upon those traits. The benefit of utilizing a pre-analyzer is that it can discriminate between the types of information retained and rejected.

This thesis demonstrates the benefits of one and two trait analyses.

3. Sorting Enhanced Information Flow Reduction Systems

An additional feature of a pre-analyzer based flow reduction system is a sort system. A sort system allows for the further refinement of information into prioritized levels. A discriminating pre-analyzer based system can place the more vital retained information higher in priority than less vital information. The decision maker then accesses the high priority items sooner than the low priority items and achieves a greater efficiency in quality information throughput.

This thesis demonstrates the benefits of adding a sort system to a pre-analyzer based flow reduction system.

B. TERMINOLOGY

The following terminology applies throughout the remainder of this thesis.

Good	when applied to information, it implies the item is important to the decision maker.
Bad	when applied to information, it implies the item is not important to the decision maker.
Value	the state of being either good or bad. No further distinction is made on the relative importance of one good (bad) item to another good (bad) item.
Trait	A feature which is used to describe an item. Traits depend upon the information under analysis but could include length, frequency, altitude, speed, origin, etc..
Trait Level	A measurement returned by a pre-analyzer which describes a trait of the item being analyzed.
Selecting	allowing an information item to proceed further through the system.
Trashing	removing or rejecting an information item from the system. Once trashed an item is unrecoverable.
Selection	the entire process of selecting and trashing information items.
Type I Error	trashing a good information item.
Type II Error	selecting a bad information item.
Stack	location of information items just prior to the decision maker analyzing them but following the flow reduction system; i.e. the decision maker's in-box.

Prioritized Stack	a stack with various levels. Each level denoting a ranking, high to low, for the decision maker to obtain information items.

C. PARAMETERS IN THE ANALYSIS

The following parameters are features of the system and the information environment and do not change.

P_G	$\equiv$	Probability an item collected from the environment is Good.
P_B	$\equiv$	Probability an item collected from the environment is Bad
	$=$	$1 - P_G$.
$\lambda_{COLLECT}$	$\equiv$	Arrival rate of items from the collection system.
μ_{DM}	$\equiv$	Service rate of the decision maker when analyzing information (Items / Unit Time).
ρ_{DM}	$\equiv$	Fraction of the time the decision maker is analyzing information (set by the decision maker).
λ_{DM}	$\equiv$	Total flow rate of items to the decision maker following flow reduction
	$=$	$\rho_{DM} * \mu_{DM}$.

These parameters provide an understanding of the information environment, the capabilities of the collection system to collect from that environment and the abilities for the decision maker to process the available information. P_G is the fraction of good information collected by the collection system. It is a function of the overall availability of good information as well as the biases incorporated in the collection system to collect good information. P_B is the fraction of information collected which is not important to the decision maker. $\lambda_{COLLECT}$ is the rate at which the collection system operates and is assumed to be much higher than the rate at which the decision maker

can process the flow, μ_{DM}. Because of this and because the decision maker must leave time to act upon the information, he/she sets a limited amount of time aside to analyze information and uses the remaining time to make decisions. This implies there is an acceptable flow rate to the decision maker, λ_{DM}, at which the decision maker can analyze all available information in his/her allotted time and then make decisions in the remaining time.

D. MEASURES OF EFFECTIVENESS

The models' measures of effectiveness [MOEs] include:

$P_{DM\ GOOD}$ ≡ Probability an important item which has been collected reaches the decision maker.

$P_{TYPE\ I\ ERROR}$ ≡ Probability an important item which has been collected does not reach the decision maker.

$P_{TRASH\ BAD}$ ≡ Probability a non-important item which has been collected does not reach the decision maker.

$P_{TYPE\ II\ ERROR}$ ≡ Probability a non-important item which has been collected reaches the decision maker.

$\rho_{DM\ GOOD}$ ≡ Fraction of the decision maker's analysis time spent analyzing important items.

$\rho_{DM\ BAD}$ ≡ Fraction of the decision maker's analysis time spent analyzing non-important items.

Define the waiting time in the system to be the delay incurred by passing an item through the flow reduction system and the decision maker. This delay begins when an item leaves the collection system and ends when the decision maker has analyzed the item or the item has been rejected by the system.

$W_{DM\ GOOD}$ ≡ Expected waiting time in the system for important items that reach the decision maker.

$W_{TYPE\ II\ ERROR} \equiv$ Expected waiting time in the system for non-important items that reach the decision maker.

$W_{TRASH\ BAD} \equiv$ Expected waiting time in the system for non-important items that are rejected by the system.

$W_{TYPE\ I\ ERROR} \equiv$ Expected waiting time in the system for important items that are rejected by the system.

These MOEs provide a number of ways of measuring the ability of the system to perform its function of flow reduction and quality enhancement. When compared across system types, $P_{DM\ GOOD}$, $P_{TYPE\ II\ ERROR}$, $P_{TYPE\ I\ ERROR}$, and $P_{TRASH\ BAD}$ demonstrate the system enhancement's ability to improve the quality of information flow to the decision maker.

The decision maker first sets his/her ρ_{DM}, the fraction of the decision maker's total time that is used to analyze information. Presumably, the decision maker would like to use much of this available time analyzing good items. Recall that $\rho_{DM\ BAD}$ is the fraction of time wasted on examining valueless information. When compared across system types it is desired that system enhancements provide an increase in $\rho_{DM\ GOOD}$ and a corresponding decrease in $\rho_{DM\ BAD}$.

The MOEs describing the expected waiting time in the system provide an understanding of the amount of time required to achieve the quality and efficiency MOEs previously described. $W_{DM\ GOOD}$ indicates the expected waiting time between the time an item is collected and the time the decision maker can act upon that piece of information. $W_{TYPE\ II\ ERROR}$ indicates the waiting time for non-important items that reach the decision maker. Because all items which reach the decision maker will eventually be processed, it is desired that the delay in the system for important items be as low as possible, and lower than that of non-important items. Conceptually, there is no loss to the decision maker if he/she never gets around to examining a non-important item but there can be a great loss in delaying an important piece of information. $W_{TRASH\ BAD}$ indicates how long the pre-processing system requires

before it removes a bad item from the flow to the decision maker. $W_{TYPE\ I\ ERROR}$ provides an indication of how long the pre-processing system requires before it errs and removes a good item from the flow to the decision maker. Since all trashing takes place prior to reaching the analytic section, these MOEs provide an understanding of the flow reduction system's efficiency in performing its function.

III. MODEL STRUCTURE

A. ARRIVAL PROCESS AND SERVICE TIME DISTRIBUTIONS

Throughout the analysis of these models the precise methods by which the devices in the system achieve their requirements are of little import. Where applicable, methods are discussed to give the reader a fundamental understanding of the issues involved. More importantly, the models provide an understanding of the implications of adding additional capabilities to information flow reduction systems. Simplifying assumptions are made to enhance that understanding. In particular, the arrival process of information from the collection system is considered to have the following features as displayed in Figure 2:

- Poisson departure process from the collection system
- Arrival rate from the collection system = $\lambda_{COLLECT}$.

Similarly, except as noted, each successive device *i* in the reduction system services each information item according to:

- Exponential service time distribution
- Service rate for device *i* = μ_i.

Accordingly, each system device sees a Poisson arrival process in steady state [Ref. 1] and uses an exponential service time distribution to achieve its requirements. Together these assumptions provide a mathematically tractable basis to conduct the analysis of the flow reduction system itself.

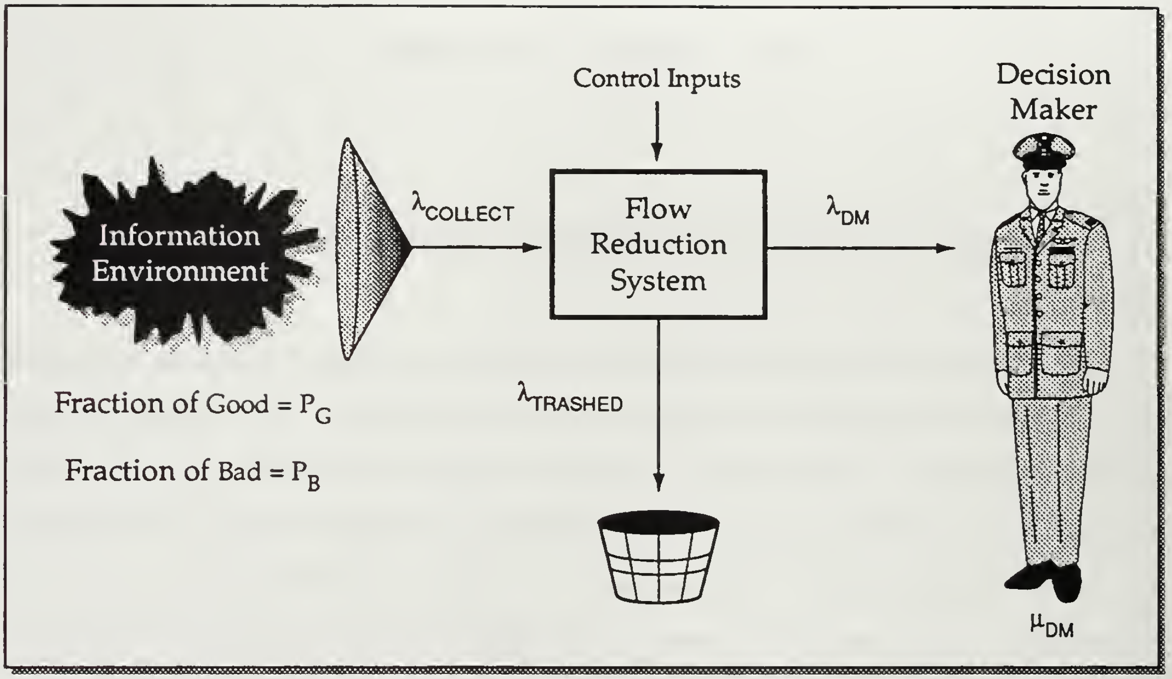

Figure 2. Flow Reduction System Overview

B. SYSTEM WITH NO FLOW REDUCTION

Information overload occurs when the decision maker is presented with more information than he/she has the ability to handle. The decision maker may attempt to examine all of the information but the rate of flow to the decision maker is greater than his/her capacity for that flow.

Define:

$\rho_{COLLECT}$ $\equiv$ Ratio of arrival rate of information from the collection system to the service rate of the decision maker when analyzing the information.

$= \lambda_{COLLECT} / \mu_{DM}$.

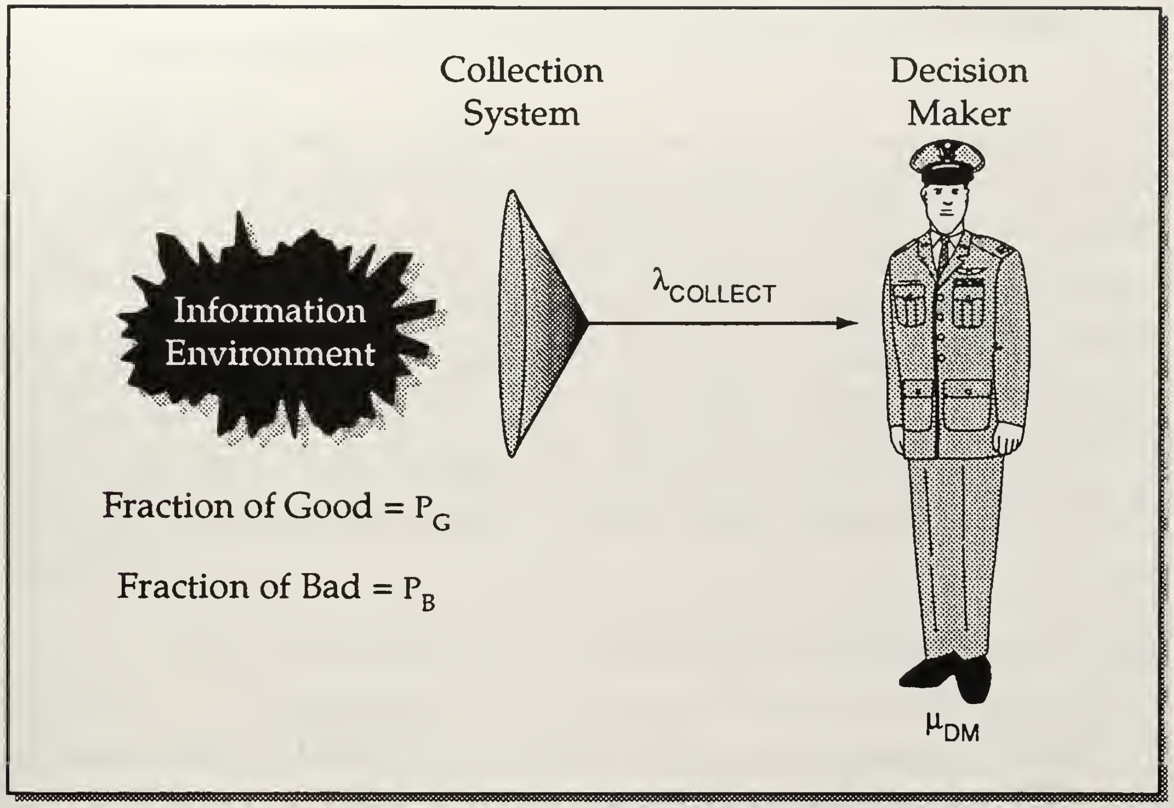

Figure 3. Simple System without Flow Reduction

If $\lambda_{COLLECT}$ is greater than or equal to μ_{DM} information is arriving faster than the decision maker can process it. The items which have been collected pile up and the decision maker is not able to process them all. Also the decision maker is spending all of his/her time analyzing information and no time acting. A flow reduction system is required to control the arrival process and reduce the fraction of time the decision maker is busy analyzing information.

The decision maker is perhaps the person best suited to judge the fraction of time, ρ_{DM}, he/she should devote to analyzing information. The remaining time can be spent acting on that information. Since ρ_{DM} is fixed by the decision maker and μ_{DM} is considered to be a fixed value for the decision maker, a flow reduction system is required if $\rho_{COLLECT}$, defined to be $\lambda_{COLLECT}$ / μ_{DM}, is greater than or equal to 1.0. A flow reduction system is also desired if $\rho_{COLLECT}$ is greater than ρ_{DM} but less than 1.0.

C. ASSUMPTIONS IN MODEL DEVELOPMENT

The following assumptions apply throughout model development:

1. The decision maker has perfect understanding of the value of an item he/she has just analyzed. This means the decision maker cannot make Type I and Type II errors.
2. The decision maker has historical information regarding the information environment and has been able to estimate its parameters. This implies the decision maker has an estimate for the probability that any item collected from the environment is Good or Bad. The decision maker also has an estimate on the distribution of trait levels of both Good and Bad items.
3. The information environment remains constant throughout the duration of analysis. There is no change to the parameters of the items collected such as:
   - probability an item collected by the collection system is good or bad,
   - rate at which items arrive from the collection system,
   - or probability that a particular good or bad item has a particular trait level.
4. No distinction is made in the service time requirements of a Good and a Bad item. Both items have the same service time distribution given they are being analyzed by the same device.
5. All devices have room for all items arriving to that device, there is no loss due to lack of capacity.
6. All devices service the arriving items on a First-Come, First-Served basis.
7. Mechanical device i services items at rate μ_i such that $\mu_i > \lambda_{COLLECT}$.
8. $\rho_{COLLECT}$ is greater than ρ_{DM}.

D. NONDISCRIMINATING FLOW REDUCTION MODEL

The first flow reduction model examined is of a system which provides a reduction of information to the decision maker without discriminating between the values of the information it is accepting or rejecting. Nor is a prioritizing system available to improve efficiency in the throughput of good information. The decision maker sees only the flow necessary to achieve ρ_{DM}. Figure 4 depicts such a process. This model produces base-case results by which enhanced models, which distinguish good items from bad, can be evaluated.

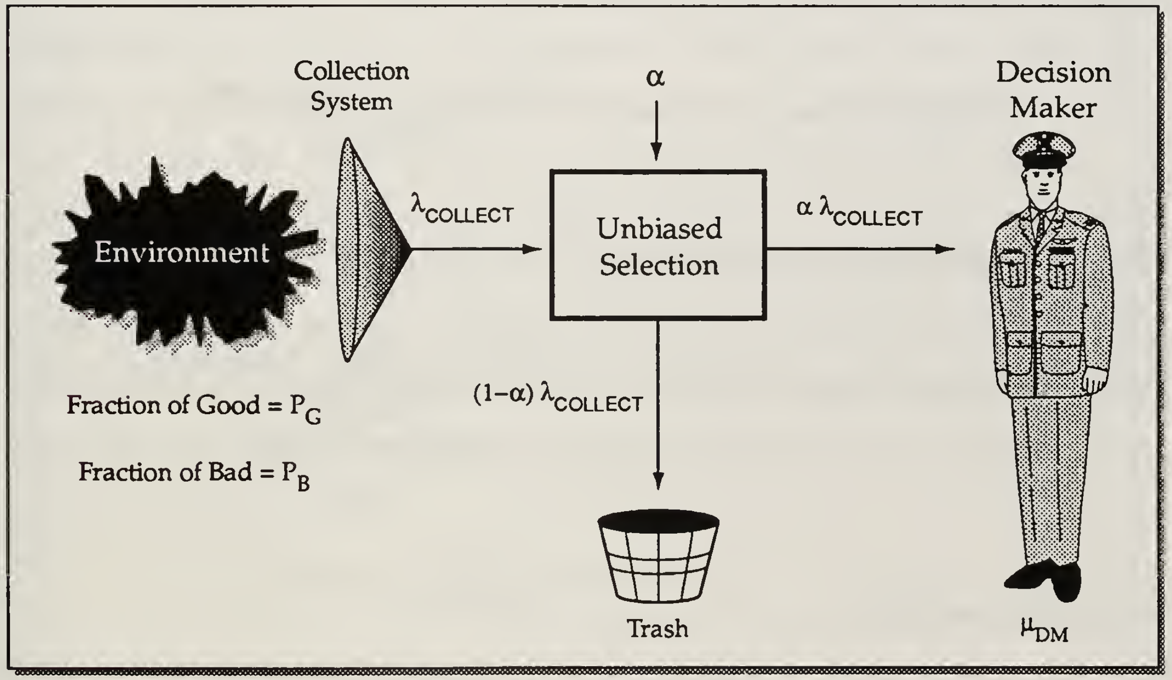

Figure 4. Nondiscriminating Flow Reduction System

Allow α to be the fraction of the collected flow that is desired to reach the decision maker. Clearly,

$$\alpha = \frac{\lambda_{DM}}{\lambda_{COLLECT}} = \frac{\rho_{DM}}{\rho_{COLLECT}}.$$

Define a success as an item surviving the selection process and a failure as trashing an item. Perform the selection process independent of any trait of the object or any previous selection or trashing and set the probability of success to α. Therefore for each item the selection process is an independent Bernoulli trial with success probability $\rho_{DM} / \rho_{COLLECT}$.

The formulas for determining the measures of effectiveness for all models are derived in the Appendix. For an understanding of the process, some of the derivations for this simple model are presented here. The probability a good item will be selected is

$$P\{\text{Value} = \text{Good} \mid \text{Selected}\} = \frac{P\{\text{Value} = \text{Good, Selected}\}}{P\{\text{Selected}\}},$$

which is simplified by independence of the selection process to

$$\begin{aligned} P\{\text{Value} = \text{Good} \mid \text{Selected}\} &= \frac{P\{\text{Value} = \text{Good}\} \cdot P\{\text{Selected}\}}{P\{\text{Selected}\}} \\ &= P\{\text{Value} = \text{Good}\} \\ &= P_G. \end{aligned}$$

In a similar fashion,

$$P\{\text{Value} = \text{Bad} \mid \text{Selected}\} = P_B,$$
$$P\{\text{Value} = \text{Good} \mid \text{Trashed}\} = P_G,$$
$$P\{\text{Value} = \text{Bad} \mid \text{Trashed}\} = P_B.$$

The probability of a Type I Error is the probability an item is trashed and it is a good item. It is also by definition the fraction of the total flow that

is good and is trashed. The probability of a Type II Error is the probability an item is selected and it is a bad item; this is also the fraction of the total flow that is bad and is selected. Thus

$$P\{\text{Type I Error}\} = P\{\text{Value} = \text{Good} \mid \text{Trashed}\} \cdot P\{\text{Trashed}\} = P_G \cdot \left(1 - \frac{\rho_{DM}}{\rho_{COLLECT}}\right),$$

$$P\{\text{Type II Error}\} = P\{\text{Value} = \text{Bad} \mid \text{Selected}\} \cdot P\{\text{Selected}\} = P_B \cdot \left(\frac{\rho_{DM}}{\rho_{COLLECT}}\right).$$

This simple system is nondiscriminating in its selection process and serves as a yardstick for measuring systems that discriminate by selecting good items over bad. This means the nondiscriminating system's Type I and Type II Error probabilities are the highest of any calibrated selection system.

Because the selection process is nondiscriminating and requires no machine effort to examine each item, we assume the time to achieve selection per item is a fixed value D. Further we assume that $D << 1/\mu_{DM}$ due to the simplicity of the task. Therefore the delay by the selector is considered to be negligible compared to the rest of the system and is ignored. This results in an exponential departure distribution from the selector with rate

$$\lambda_{DEPARTURE} = \frac{\rho_{DM}}{\rho_{COLLECT}} \cdot \lambda_{COLLECT} = \lambda_{DM} = \alpha \cdot \lambda_{COLLECT}$$

which corresponds to a Poisson arrival process to the decision maker. Because the decision maker's service time distribution is exponential, this system is a single exponential channel queue (M/M/1). If an item arrives to the decision maker when the decision maker is idle (not analyzing another item or making a decision), the decision maker will immediately begin analyzing the item. If the item arrives when the decision maker is busy the item is placed in a queue and will be analyzed by the decision maker on a

First-Come, First-Served basis. Waiting time measures of effectiveness are easily obtained from an M/M/1 queueing system, [Ref. 1], by the equation

$$W = \begin{cases} \dfrac{1}{\left(\mu_{\text{SERVICE RATE}} - \lambda_{\text{ARRIVALRATE}}\right)}, & \mu_{\text{SERVICE RATE}} > \lambda_{\text{ARRIVALRATE}} \\ \infty, & \text{else.} \end{cases}$$

The measures of effectiveness for this model are displayed in Table 2 through Table 4 with numerical results obtained using the example data provided in Table 1.

EXAMPLE DATA FOR EVALUATION OF MOEs		
Parameter	**Equation**	**Value**
Probability an item from the environment is good.	P_G	0.3
Probability an item from the environment is bad.	$P_B = (1 - P_G)$	0.7
Service rate of the decision maker.	μ_{DM}	10 Items/Hr
Arrival rate from the collection system	$\lambda_{COLLECT}$	30 Items/Hr
Ratio of the collection rate to the decision maker's analysis rate. Impossible to achieve all analysis if greater than 1.0.	$\rho_{COLLECT} = \dfrac{\lambda_{COLLECT}}{\mu_{DM}}$	3.0
Fraction of the decision maker's time the decision maker spends analyzing.	ρ_{DM}	0.8

Table 1. Nondiscriminating Flow Reduction Model Example

PROBABILITY MEASURES OF EFFECTIVENESS		
MOE	**Equation**	**Example Result**
$P_{DM\ GOOD}$ Probability an item is good and it reaches the decision maker.	$P_G \cdot \left(\frac{\rho_{DM}}{\rho_{COLLECT}}\right)$	0.080
$P_{TYPE\ I\ ERROR}$ Probability an item is good and it is trashed.	$P_G \cdot \left(1 - \frac{\rho_{DM}}{\rho_{COLLECT}}\right)$	0.220
$P_{TRASH\ BAD}$ Probability an item is bad and it is trashed.	$P_B \cdot \left(1 - \frac{\rho_{DM}}{\rho_{COLLECT}}\right)$	0.513
$P_{TYPE\ II\ ERROR}$ Probability an item is bad and it reaches the decision maker.	$P_B \cdot \left(\frac{\rho_{DM}}{\rho_{COLLECT}}\right)$	0.187
Total	**1.00**	**1.00**

Table 2. Nondiscriminating Flow Reduction Model Probability MOEs

ANALYSIS TIME MEASURES OF EFFECTIVENESS		
MOE	**Equation**	**Example Result**
$\rho_{DM\ GOOD}$ Fraction of decision maker's analysis time spent on good items	$\frac{P_{DM\ GOOD}}{P_{DM\ GOOD} + P_{TYPE\ II\ ERROR}}$	0.30
$\rho_{TYPE\ II\ ERROR}$ Fraction of decision maker's analysis time spent on bad items	$\frac{P_{TYPE\ II\ ERROR}}{P_{DM\ GOOD} + P_{TYPE\ II\ ERROR}}$	0.70

Table 3. Nondiscriminating Flow Reduction Model Analysis Time MOEs

WAITING TIME MEASURES OF EFFECTIVENESS		
MOE	**Equation**	**Example Result**
$W_{\text{DM GOOD}}$ Expected total time in the system for good items reaching the decision maker.	$\dfrac{1}{\mu_{DM} - \lambda_{DM}}$	0.50 Hrs.
$W_{\text{TYPE II ERROR}}$ Expected total time in the system for bad items reaching the decision maker.	$\dfrac{1}{\mu_{DM} - \lambda_{DM}}$	0.50 Hrs.
$W_{\text{TRASH BAD}}$ Expected total time in the system for bad items which are trashed.	0.0	0.0
$W_{\text{TYPE I ERROR}}$ Expected total time in the system for good items which are trashed.	0.0	0.0

Table 4. Nondiscriminating Flow Reduction Model Waiting Time MOEs

E. DISCRIMINATING FLOW REDUCTION MODEL - SINGLE TRAIT

A trait-based selection process seeks to improve the quality of the information reaching the decision maker while still reducing the overall flow from $\lambda_{COLLECT}$ to λ_{DM}. To achieve the enhanced quality, the selectors discriminate to select items that have traits known to be more prevalent in good items. The result is a probabilistic technique for selection. Woosley, [Ref. 2], provides a description of a number of techniques which achieve this goal. The methods include: classical inference, Baysian inference, Dempster/Shafer inference, fuzzy set theory, cluster analysis, estimation theory and entropy. To utilize this method the decision maker must possess a sensor which has the ability to examine each item and determine the trait level associated with a particular trait of that item. The decision maker must also believe that the trait being analyzed will assist him in selecting good items over bad items.

For example, the decision maker may be receiving information in the form of messages. Assume that historically, messages which are good tend to be longer than messages which are bad. The decision maker chooses to use a sensor which is capable of measuring the length of each message. The selector examines the result of that measurement, selects the longer messages and trashes the shorter ones. Note that the trait "message length" is common to all information items of the type "message". We require each item to possess the trait the sensor is examining.

To prepare for selection based upon trait, the decision maker examines historical records of information items of this type and separates them into two groups: items which are good and items which are bad. The sensor is used to examine the trait level of each item and probability densities for both the good items, $f_{GOOD}(t)$, and the bad items, $f_{BAD}(t)$, over the trait levels are formed as demonstrated in Figure 5. Note the two densities overlap but display a separation which assists in differentiating good items from bad.

Note also that there is no requirement for a good (bad) item to be present at all trait levels, nor is there a requirement for the densities $f_{BAD}(t)$ and $f_{GOOD}(t)$ to be of the same type. In fact, the further the two densities are separated by trait level and skewness the more useful the trait is for use in the selection process. Completely separate densities would provide perfect classification of items.

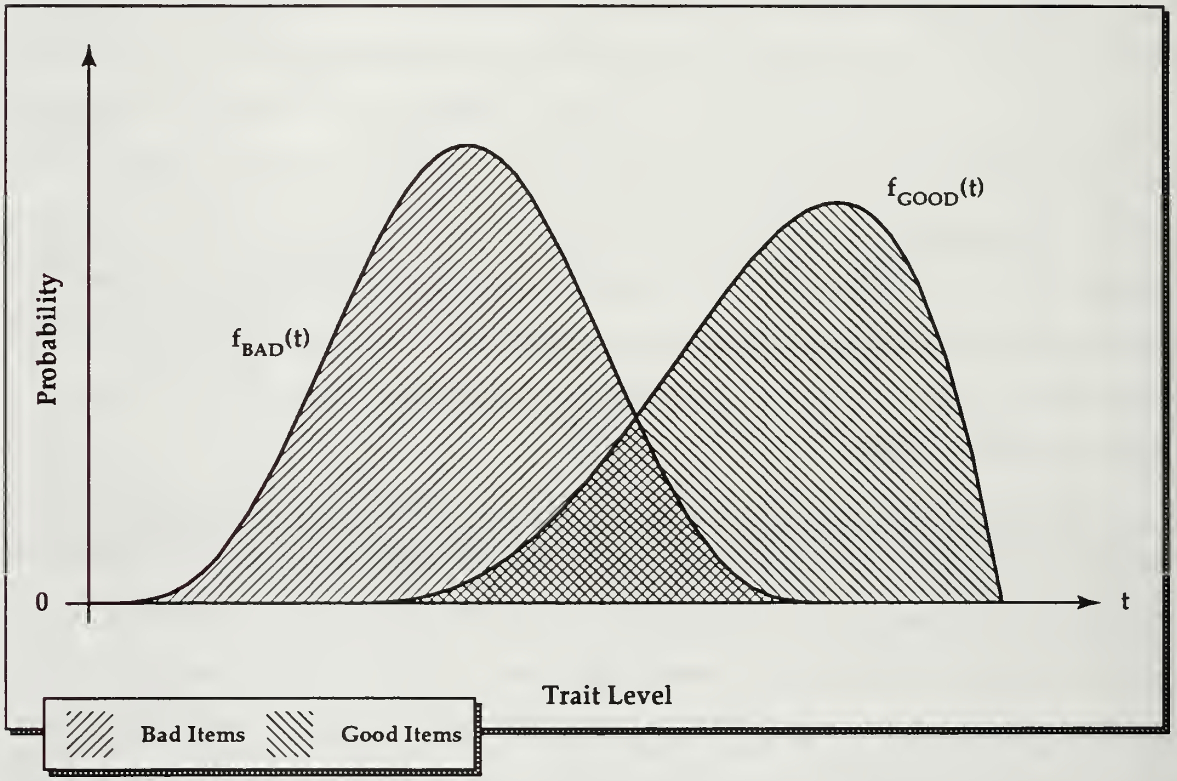

Figure 5. Trait Level Probability Densities By Item Value

Define

F	$\equiv$	Sensor for analyzing the single trait used for selection.
μ^F	$\equiv$	Rate at which F determines trait levels for each item. The rate μ^F is independent of the value of the item.
ρ^F	$\equiv$	Fraction of the time F is analyzing information.

As in the nondiscriminating case, the time required to compare the measurement result from the sensor with known selection cutoff criteria is considered to be negligible when compared with the sensor's processing time

and the decision maker's analysis time. It is therefore ignored in this model. Because F is a machine and capable of much higher speeds than the human decision maker, $\mu^F > \mu_{DM}$, and by Assumption 7 (section C) $\mu^F > \lambda_{COLLECT}$ or equivalently $0 < \rho^F < 1.0$.

This separation process provides three distinct areas of interest defined by the parameters of $f_{GOOD}(t)$ and $f_{BAD}(t)$ as demonstrated in Figure 6. From trait level T_0 to trait level T_1, Area I, the probability that an item is good is zero and the probability that an item is bad is positive. Therefore, items displaying a trait level $t \in [T_0, T_1)$ should be trashed since they are bad with probability 1.0. Area III shows the probability that an item is good is positive while the probability that an item is bad zero. Therefore, any items displaying a trait level $t \in [T_2, T_3]$ should always be selected since they are good with probability 1.0. Areas I and II take advantage of the offset in the distributions. Area II, $[T_1, T_2)$, has positive probabilities for both good and bad items. The decision maker must have a method to determine which items from Area II should be selected and which trashed. Here the advantages of skewed distributions becomes apparent. In this case it is obvious that choosing an item with a higher trait level will result in a greater probability that the item is good. Let $S^F \in [T_1, T_2)$ be the trait level above which the selector will select the item and below which the selector will trash the item. A method for determining the appropriate S^F is discussed in Chapter IV. The resultant system is shown in Figure 7.

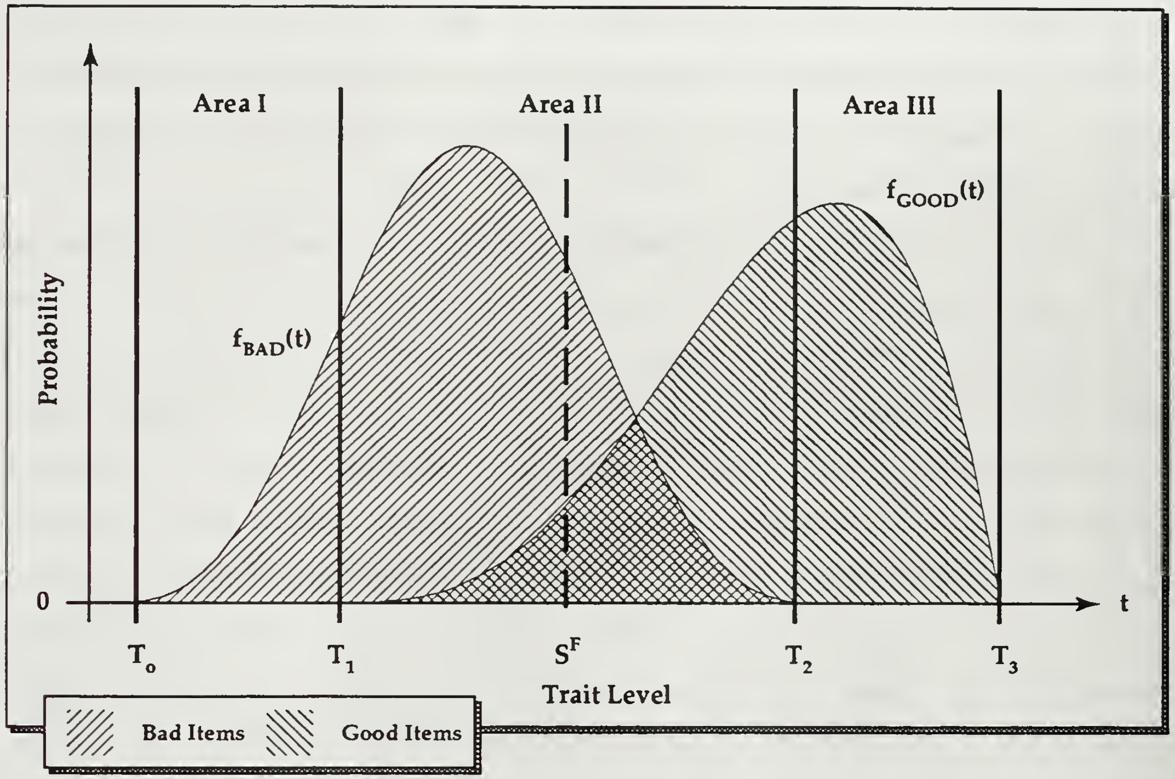

Figure 6. Areas of Interest for Trait Separation Densities

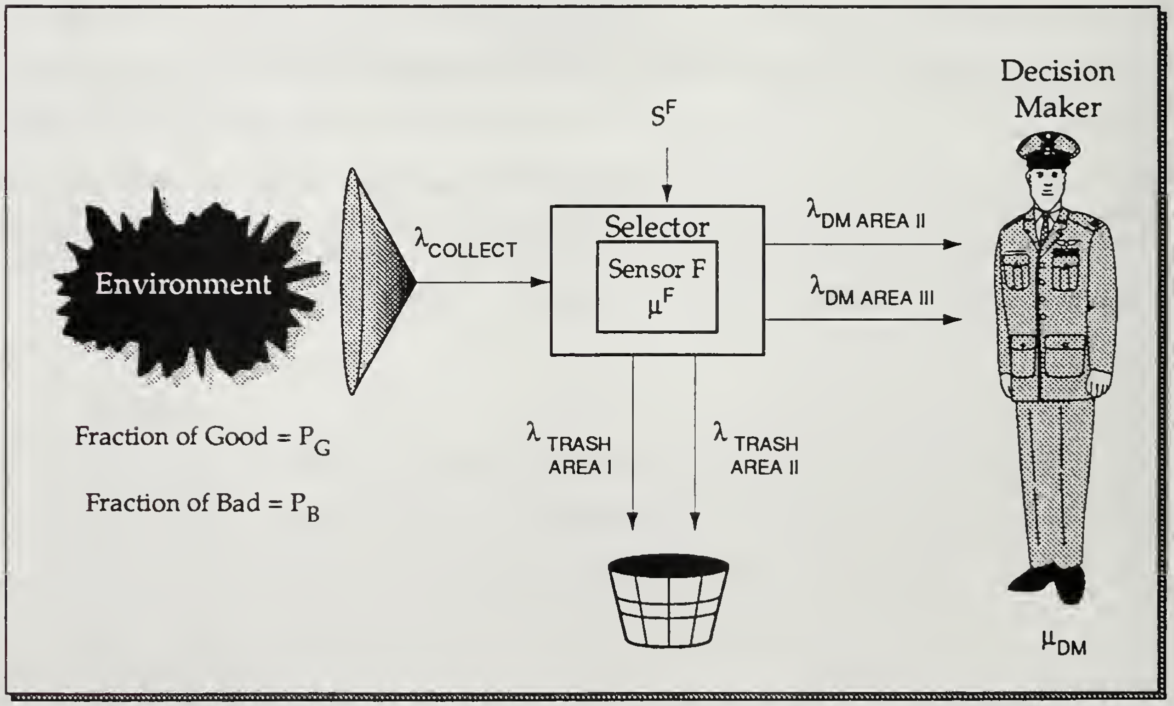

Figure 7. Single Trait Discriminating Flow Reduction System

Appendix A contains the derivations of the measures of effectiveness. Table 5 provides example data for MOE evaluation. The resultant measures of effectiveness are collected in Table 6 through Table 8.

EXAMPLE DATA FOR EVALUATION OF MOEs		
Parameter	Equation	Value
Probability an item from the environment is good.	P_G	0.3
Probability an item from the environment is bad.	P_B	0.7
Service rate of the decision maker.	μ_{DM}	10 Items/Hr
Arrival Rate from the collection system.	$\lambda_{COLLECT}$	30 Items/Hr
Ratio of the collection rate to the decision maker's analysis rate. Impossible to achieve all analysis if greater than 1.0.	$\rho_{COLLECT}$	3.0
Fraction of the decision maker's time the decision maker spends analyzing.	ρ_{DM}	0.8
Lowest trait level for bad items.	T_0	0.0
Lowest trait level for good items.	T_1	0.1
Highest trait level for bad items.	T_2	1.0
Highest trait level for good items.	T_3	1.1
Cutoff trait level. S^F set to reduce flow to λ_{DM}	S^F	0.701
Probability densities : good items.	$f_{GOOD}(t)$	*Beta*(t; 4, 2, 0.1, 1.1)
Probability densities : bad items.	$f_{BAD}(t)$	*Beta*(t; 5, 5, 0.0, 1.0)
Service rate for sensor F.	μ^F	50 Items/Hr.

Table 5. Single Trait Flow Reduction Model Example

Recall that the beta density on the interval (A, B) is given by [Ref. 3]

$$Beta(t;\ \alpha,\ \beta,\ A,B)=\begin{cases}\dfrac{1}{B-A}\cdot\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)}\cdot\left(\dfrac{x-A}{B-A}\right)^{(\alpha-1)}\cdot\left(\dfrac{B-x}{B-A}\right)^{(\beta-1)} & A\le x\le B\\ 0 & \text{else}\end{cases}.$$

Define $F_{GOOD}(t)$ and $F_{BAD}(t)$ as the cumulative distribution functions (CDFs) of $f_{GOOD}(t)$ and $f_{BAD}(t)$ respectively.

PROBABILITY MEASURES OF EFFECTIVENESS		
MOE	**Equation**	**Example Result**
$P_{DM\ GOOD}$ Probability an item is good and it reaches the decision maker.	$P_G\cdot\left(1-F_{GOOD}\left(S^F\right)\right)$	0.198 *(Base-Case = 0.080)*
$P_{TYPE\ I\ ERROR}$ Probability an item is good and it is trashed.	$P_G\cdot F_{GOOD}\left(S^F\right)$	0.102 *(Base-Case = 0.220)*
$P_{TRASH\ BAD}$ Probability an item is bad and it is trashed.	$P_B\cdot F_{BAD}\left(S^F\right)$	0.632 *(Base-Case = 0.513)*
$P_{TYPE\ II\ ERROR}$ Probability an item is bad and it reaches the decision maker..	$P_B\cdot\left(1-F_{BAD}\left(S^F\right)\right)$	0.068 *(Base-Case = 0.187)*
Total	**1.00**	**1.00**

Table 6. Single Trait Flow Reduction Model Probability MOEs

ANALYSIS TIME MEASURES OF EFFECTIVENESS		
MOE	Equation	Example Result
$\rho_{DM\ GOOD}$ Fraction of the decision maker's analysis time spent on good items.	$\frac{P_{DM\ GOOD}}{P_{DM\ GOOD} + P_{TYPE\ II\ ERROR}}$	0.74 *(Base-Case = 0.30)*
$\rho_{TYPE\ II\ ERROR}$ Fraction of the decision maker's analysis time spent on bad items.	$\frac{P_{TYPE\ II\ ERROR}}{P_{DM\ GOOD} + P_{TYPE\ II\ ERROR}}$	0.26 *(Base-Case = 0.70)*

Table 7. Single Trait Flow Reduction Model Analysis Time MOEs

WAITING TIME MEASURES OF EFFECTIVENESS		
MOE	Equation	Example Result
$W_{\text{DM GOOD}}$ Expected total time in the system for good items reaching the decision maker.	$\dfrac{1}{\mu^F - \lambda_{\text{COLLECT}}} + \dfrac{1}{\mu_{\text{DM}} - \lambda_{\text{DM}}}$	0.55 Hrs. *(Base-Case = 0.50)*
$W_{\text{TYPE II ERROR}}$ Expected total time in the system for bad items reaching the decision maker.	$\dfrac{1}{\mu^F - \lambda_{\text{COLLECT}}} + \dfrac{1}{\mu_{\text{DM}} - \lambda_{\text{DM}}}$	0.55 Hrs. *(Base-Case = 0.50)*
$W_{\text{TRASH BAD}}$ Expected total time in the system for bad items which are trashed.	$\dfrac{1}{\mu^F - \lambda_{\text{COLLECT}}}$	0.05 Hrs. *(Base-Case = 0.0)*
$W_{\text{TYPE I ERROR}}$ Expected total time in the system for good items which are trashed.	$\dfrac{1}{\mu^F - \lambda_{\text{COLLECT}}}$	0.05 Hrs. *(Base-Case = 0.0)*

Table 8. Single Trait Flow Reduction Model Waiting Time MOEs

The results from the example data, when compared with those from the base-case, demonstrate the power and cost associated with discriminating selection. With a single trait system the probabilities of retaining a good and trashing a bad item are improved by 150 and over 20 percent respectively while the probabilities of Type I and Type II errors decreased to 46 percent and 36 percent of their original values. The fraction of the time the analyst is working on good information improved by 150 percent and the fraction of

time spent on bad items decreased to about a third of its original value. These improvements are not without a cost. The waiting time MOEs demonstrate that the trait based selection process induces a delay in the arrival of information. That delay is seen by all items entering the system and, in this instance, increases the system delay by 0.05 hours or a ten percent increase for selected items. It is therefore important for the decision maker to understand the implications of a discriminating selection enhancement when designing a flow reduction system. The sensitivity of the system to minor changes is discussed in Chapter IV.

F. DISCRIMINATING FLOW REDUCTION MODEL - DUAL TRAIT

A multiple trait flow reduction system is similar to the single trait system in its discriminating towards selecting good items over bad and its capability for reducing overall flow to the decision maker. These systems use a network of selectors to perform the same task as a single selector system but with the goal of further enhancing the quality of the information reaching the decision maker and reducing the Type I and Type II errors the system produces.

As illustrated in Figure 8, the network of interest is two selectors in tandem. Other selector networks can be formed by extending this model.

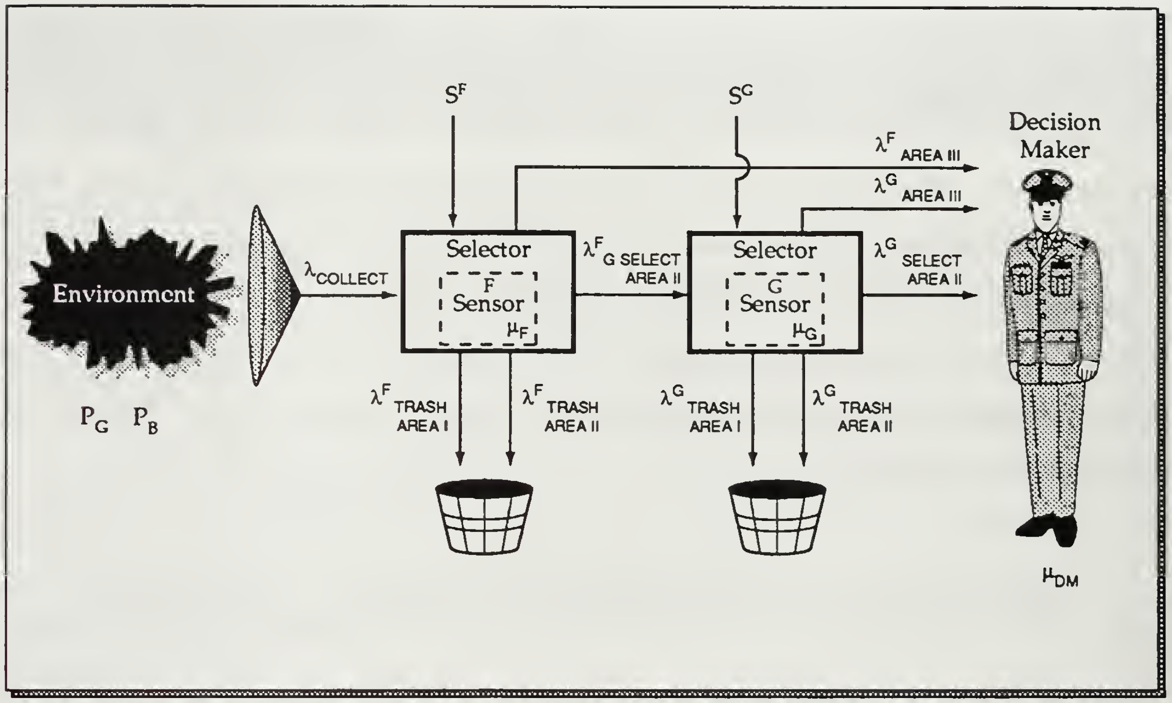

Figure 8. Two Trait Flow Reduction System

To perform multiple trait analysis we require a different sensor to measure each trait. Define:

F	Sensor associated with the first selector,
μ^F	Service rate for sensor F,
ρ^F	Fraction of time the sensor F is analyzing an item,
G	Sensor associated with the second selector,
μ^G	Service rate for sensor G,
ρ^G	Fraction of time the sensor G is analyzing an item.

Again, we require $\mu^i > \lambda_{COLLECT}$ and $0 < \rho^i < 1$, for all $i \ \varepsilon \ \{F, G\}$.

We also seek traits which exhibit conditional independence. This implies that knowing the value of an item and its trait level for the first trait does not help in predicting the trait level of the second trait or,

$$P\{F = x \mid \text{Value} = v\} \cdot P\{G = y \mid \text{Value} = v\} = P\{F = x, G = y \mid \text{Value} = v\} \quad \forall \ x, y, v.$$

Therefore a reduction in the flow by any selector *i* preceding selector *j* does not change the probability distributions associated with trait *j*.

Figure 9 depicts the two sets of trait level probability distributions as sampled by sensor F and sensor G respectively. The parameters are labeled similarly to the single trait model with the addition of a superscript to indicate the sensor which is measuring the item. Both traits display a separation across their respective trait levels, {F Area I, F Area III, G Area I, G Area III}, which can be exploited by the selection process. The probability density functions (PDFs) associated with the first trait are $f_{GOOD}(t)$ and $f_{BAD}(t)$ for good and bad items respectively. Their CDFs are $F_{GOOD}(t)$ and $F_{BAD}(t)$. Similarly for the second trait the PDFs and CDFs are $g_{GOOD}(t)$, $g_{BAD}(t)$, $G_{GOOD}(t)$ and $G_{BAD}(t)$. The cutoff trait level used by the first selector is S^F and the cutoff trait level for the second selector is S^G.

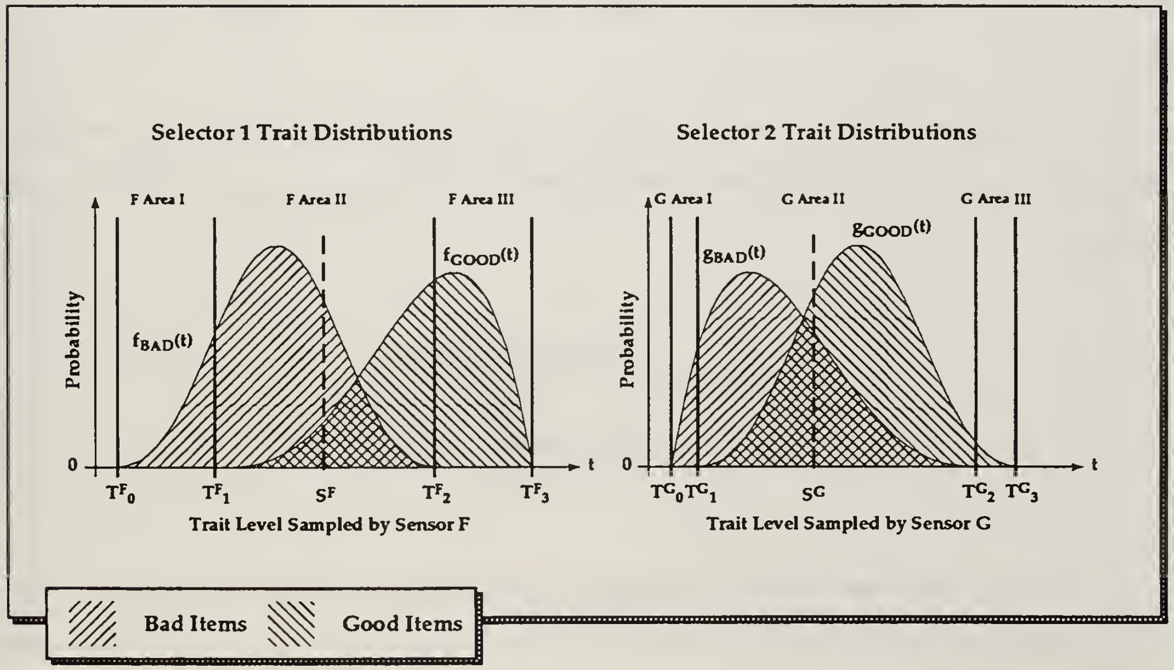

Figure 9. Two Trait Selection Model Probability Densities by Trait Level

Sensor F is associated with the first selection process and performs its function in the same manner as the sensor in the single trait model. The output from Area I of sensor F's densities plot, Figure 9, as well as that in F

Area II which is less than S^F is still sent to the trash. The output from F Area III is routed directly to the decision maker, by-passing the second selector. However, in the two trait selection model the selected items from F Area II are no longer routed directly to the decision maker but instead proceed to the second selector. The flow rate into the second selector is therefore dependent upon the values T^F_2 and S^F as well as the distributions $f_{GOOD}(t)$ and $f_{BAD}(t)$ for the first trait. This rate is calculated in Appendix A and is:

$$\lambda^F_G = \left(P_G \cdot \left[F_{GOOD}\left(T^F_2\right) - F_{GOOD}\left(S^F\right)\right] + P_B \cdot \left[F_{BAD}\left(T^F_2\right) - F_{BAD}\left(S^F\right)\right]\right) \cdot \lambda_{COLLECT} \cdot$$

To simplify follow-on equations define the probability a collected item is routed from F to G as

$$P_{F\,TO\,G} = \left(P_G \cdot \left[F_{GOOD}\left(T^F_2\right) - F_{GOOD}\left(S^F\right)\right] + P_B \cdot \left[F_{BAD}\left(T^F_2\right) - F_{BAD}\left(S^F\right)\right]\right).$$

Using this notation we calculate the probability an arrival to the second selector is good or bad. Define these probabilities as $P_{GOOD\,|\,F\,TO\,G}$ and $P_{BAD\,|\,F\,TO\,G}$ where,

$$P_{GOOD\,|\,F\,TO\,G} = \frac{P_{SELECTED\ GOOD\ F\ AREA\ II}}{P_{F\,TO\,G}},$$

$$P_{BAD\,|\,F\,TO\,G} = \frac{P_{SELECTED\ BAD\ F\ AREA\ II}}{P_{F\,TO\,G}}.$$

Except that the arrival flow rate is determined by the first selector and the probability that an arrival to the selector is good has changed, the second selector works exactly like the single trait selector previously modeled. Because the trait levels are conditionally independent given value, any adjustment in the cutoff setting, S^F, for the first selector does not impact the probability distributions for the second trait.

Sample data for the parameters of the system are found in Table 9. The measures of effectiveness for the two trait model are presented in Table 10 through Table 12.

EXAMPLE DATA FOR EVALUATION OF MOEs		
Parameter	Equation	Value
Probability an item from the environment is good.	P_G	0.3
Probability an item from the environment is bad.	P_B	0.7
Service rate of the decision maker.	μ_{DM}	10 Items / Hr.
Arrival rate from the collection system.	$\lambda_{COLLECT}$	30 Items / Hr.
Ratio of the collection rate to the decision maker's analysis rate. Impossible to achieve all analysis if greater than 1.0.	$\rho_{COLLECT}$	3.0
Fraction of the decision maker's time the decision maker spends analyzing.	ρ_{DM}	0.8
First Trait : Lowest trait level for bad items.	T^F_0	0.0
First Trait : Lowest trait level for good items.	T^F_1	0.1
First Trait : Highest trait level for bad items.	T^F_2	1.0
First Trait : Highest trait level for good items.	T^F_3	1.1
First Trait : Cutoff trait level.	S^F	0.6350
First Trait : PDF good items.	$f_{GOOD}(t)$	*Beta*(t; 4, 2, 0.1, 1.1)
First Trait : PDF bad items.	$f_{BAD}(t)$	*Beta*(t; 5, 5, 0.0, 1.0)
Service rate for sensor F	μ^F	50 Items / Hr.
Second Trait : Lowest trait level for bad items.	T^G_0	0.0
Second Trait : Lowest trait level for good items.	T^G_1	0.1
Second Trait : Highest trait level for bad items.	T^G_2	1.0
Second Trait : Highest trait level for good items.	T^G_3	1.1
Second Trait : Cutoff trait level.	S^G	0.3783
Second Trait : Probability distribution good items.	$g_{GOOD}(t)$	*Beta*(t; 5, 5, 0.1, 1.1)
Second Trait : Probability distribution bad items.	$g_{BAD}(t)$	*Beta*(t; 2, 4, 0.0, 1.0)
Service rate for sensor F	μ^G	50 Items / Hr.

Table 9. Two Trait Flow Reduction Model Example

PROBABILITY MEASURES OF EFFECTIVENESS		
MOE	Equation	Example Result
$P_{DM\ GOOD}$ Probability an item is good and it reaches the decision maker.	$\left[P_G \cdot \left(1 - F_{GOOD}\left(T_2^F\right)\right)\right]$ + $\left[P_{F\ TO\ G} \cdot P_{GOOD\ \mid\ F\ TO\ G} \cdot \left(1 - G_{GOOD}\left(S^G\right)\right)\right]$	0.214 *(Single Trait = 0.198)* *(Base-Case = 0.080)*
$P_{TYPE\ I\ ERROR}$ Probability an item is good and it is trashed.	$\left[P_G \cdot F_{GOOD}\left(S^F\right)\right]$ + $\left[P_{F\ TO\ G} \cdot P_{GOOD\ \mid\ F\ TO\ G} \cdot G_{GOOD}\left(S^G\right)\right]$	0.086 *(Single Trait = 0.102)* *(Base-Case = 0.220)*
$P_{TRASH\ BAD}$ Probability an item is bad and it is trashed	$P_B \cdot F_{BAD}\left(S^F\right)$ + $P_{F\ TO\ G} \cdot P_{BAD\ \mid\ F\ TO\ G} \cdot G_{BAD}\left(S^G\right)$	0.648 *(Single Trait = 0.632)* *(Base-Case = 0.513)*
$P_{TYPE\ II\ ERROR}$ Probability an item is bad and it reaches the decision maker.	$P_{F\ TO\ G} \cdot P_{BAD\ \mid\ F\ TO\ G} \cdot \left(1 - G_{BAD}\left(S^G\right)\right)$	0.052 *(Single Trait = 0.068)* *(Base-Case = 0.187)*
Total	1.00	1.00

Table 10. Two Trait Flow Reduction Model Probability MOEs

ANALYSIS TIME MEASURES OF EFFECTIVENESS		
MOE	Equation	Example Result
$\rho_{DM\ GOOD}$ Fraction of the decision maker's analysis time spent on good items.	$\dfrac{P_{DM\ GOOD}}{P_{DM\ GOOD} + P_{TYPE\ II\ ERROR}}$	0.805 *(Single Trait = 0.74)* *(Base-Case = 0.30)*
$\rho_{TYPE\ II\ ERROR}$ Fraction of the decision maker's analysis time spent on bad items.	$\dfrac{P_{TYPE\ II\ ERROR}}{P_{DM\ GOOD} + P_{TYPE\ II\ ERROR}}$	0.195 *(Single Trait = 0.26)* *(Base-Case = 0.70)*

Table 11. Two Trait Flow Reduction System Model Analysis Time MOEs

WAITING TIME MEASURES OF EFFECTIVENESS		
MOE	Equation	Example Result
$W_{DM\ GOOD}$ Expected total time in the system for good items reaching the decision maker.	$\frac{1}{\mu^F-\lambda_{COLLECT}}+\left(\frac{P_{FTOG}\cdot P_{GOOD\ IF\ TOG}\cdot\left(1-G_{GOOD}(S^G)\right)}{P_G\cdot\left(1-F_{GOOD}(T_2^F)\right)+\left[P_{FTOG}\cdot P_{GOOD\ IF\ TOG}\cdot\left(1-G_{GOOD}(S^G)\right)\right]}\right)\cdot\left(\frac{1}{\mu^G-\lambda_G^F}\right)+\left(\frac{1}{\mu_{DM}-\lambda_{DM}}\right)$	0.571 Hrs. *(Single Trait = 0.55)* *(Base-Case = 0.50)*
$W_{TYPE\ II\ ERROR}$ Expected total time in the system for bad items reaching the decision maker.	$\left(\frac{1}{\mu^F-\lambda_{COLLECT}}\right)+\left(\frac{1}{\mu^G-\lambda_G^F}\right)+\left(\frac{1}{\mu_{DM}-\lambda_{DM}}\right)$	0.574 Hrs. *(Single Trait = 0.55)* *(Base-Case = 0.50)*
$W_{TRASH\ BAD}$ Expected total time in the system for bad items which are trashed.	$\left(\frac{1}{\mu^F-\lambda_{COLLECT}}\right)+\left(\frac{P_{FTOG}\cdot P_{BAD\ IF\ TOG}\cdot G_{BAD}(S^G)}{\left[P_{FTOG}\cdot P_{BAD\ IF\ TOG}\cdot G_{BAD}(S^G)\right]+P_B\cdot F_{BAD}(S^F)}\right)\cdot\left(\frac{1}{\mu^G-\lambda_G^F}\right)$	0.053 Hrs. *(Single Trait = 0.05)* *(Base-Case = 0.00)*
$W_{TYPE\ I\ ERROR}$ Expected total time in the system for good items which are trashed.	$\left(\frac{1}{\mu^F-\lambda_{COLLECT}}\right)+\left(\frac{P_{FTOG}\cdot P_{GOOD\ IF\ TOG}\cdot G_{GOOD}(S^G)}{\left[P_{FTOG}\cdot P_{GOOD\ IF\ TOG}\cdot G_{GOOD}(S^G)\right]+P_G\cdot F_{GOOD}(S^F)}\right)\cdot\left(\frac{1}{\mu^G-\lambda_G^F}\right)$	0.054 Hrs. *(Single Trait = 0.05)* *(Base-Case = 0.00)*

Table 12. Two Trait Flow Reduction Model Waiting Time MOEs

G. SORTING ENHANCED FLOW REDUCTION MODEL

The single trait flow reduction model demonstrates the ability of a discriminating selection system to improve the quality of the information reaching the decision maker. The dual trait system improves the quality even further and allows a slight time savings for good items when compared with bad items in the same reduction system. This time savings is the result of items which are known to be good bypassing the second selector and proceeding directly to the decision maker's in-box. Once in the in-box, these known good items must wait in a first-come, first-served queue behind items which may be good or bad. The sorting process is designed to reduced the waiting time of good items by giving priority within that in-box to items which have a greater probability of being good. Since the priority system does not reduce the number of bad items reaching the decision maker, it will not improve the overall quality of the flow. That is achieved only by the selectors. It will, however, allow the good items to reach the decision maker sooner.

The sorting enhancement model replaces the single slot in-box used by the decision maker with a three level prioritized in-box. The model examines the benefits and costs associated with adding this in-box to a single trait flow reduction system.

Define the three slots of the in-box as 1, 2, and 3. Let slot 1 be the slot that the decision maker accesses first when reaching for a new information item to analyze. If there are no items within slot 1 the decision maker obtains an item from slot 2. Should both slot 1 and slot 2 be empty the decision maker obtains an item from the third slot. If all slots are empty the decision maker obtains the next item arriving to any slot in the in-box immediately upon its arrival. If an item arrives while the decision maker is busy, it is placed into the appropriate slot and remains there to be serviced in first-come, first-served order. No arriving item can interrupt a busy decision maker who is analyzing another item, even if the new item holds a higher priority.

The in-box and decision maker combination described is a non-preemptive single channel queueing system with a Poisson arrival process and an exponential service rate distribution. The waiting times in queue associated with each slot $i \in \{1, 2, 3\}$ of the in-box are calculated using [Ref. 4],

$$W_q^{(i)} = \frac{\sum_{k=1}^{r} \left(\frac{\lambda_k}{\mu_k^2} \right)}{(1-\sigma_{i-1})(1-\sigma_i)}$$

where

r	≡	total number of slots (3),
λ_k	≡	the arrival rate to slot $k \in \{1, 2, 3\}$,
μ_k	≡	the service rate for items of the type placed in slot k,
	=	μ_{DM} for all slots,
σ_i	≡	$\sum_{j=1}^{i} \frac{\lambda_j}{\mu_j}, \quad (\sigma_0 = 0).$

To obtain the total waiting time associated with the in-box slot as well as the decision maker, the expected service time per item, $1/\mu_{DM}$, is added to this value.

The arrival rate, λ_i, to each slot i is determined by how the selectors feed the in-box. Because the objective is to analyze good information, items will be placed by priority in the in-box with the items that have the highest probability of being good items in the highest slot. Items of lower probability will fill lower slots. From the single trait flow reduction model we know that items in Area III, Figure 10, are good with probability 1.0. Clearly, we should forward all items with trait levels in the interval $(T_2, T_3]$ to slot 1 of the priority in-box. Define S^F as the cutoff trait level for the selection process of this single selector. The remaining items that have been selected to reach the decision maker have trait levels on the interval $[S^F, T_2]$. We must define a method of dividing this remaining flow into the two remaining in-box slots. As in the single trait model, the skewness of the densities $f_{GOOD}(t)$ and $f_{BAD}(t)$ demonstrate that items with higher trait levels tend to have a greater

probability of being good items. Define S_1 as a cutoff trait level such that items on the interval $[S_1, T_2]$ are placed in slot 2 of the in-box and items on the interval $[S^F, S_1)$ are placed in slot 3, where $S^F \leq S_1 \leq T_2$. Figure 11 depicts this selection and prioritizing process.

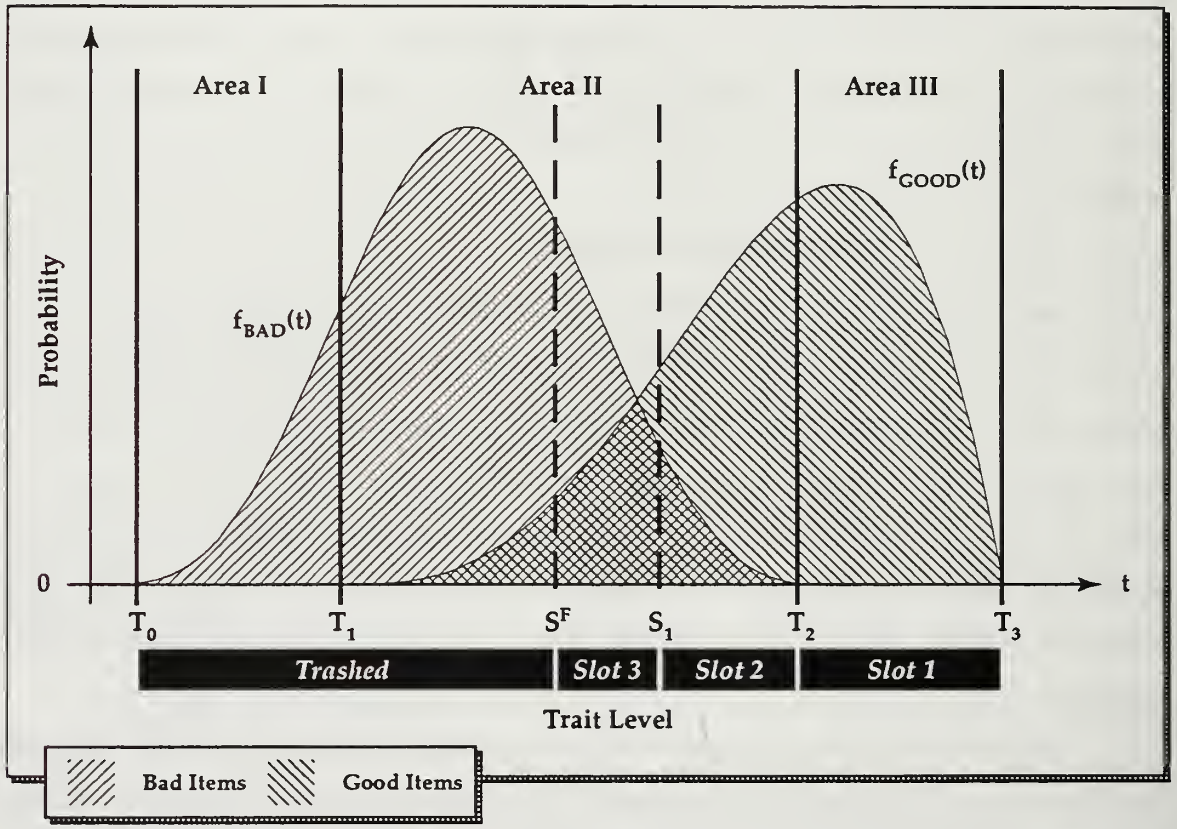

Figure 10. Single Trait Flow Reduction Model with Prioritized In-box Cutoffs

The arrival rates to each in-box slot i are:

$$\lambda_1 = P_G \cdot (1 - F_{GOOD}(T_2)) \cdot \lambda_{COLLECT},$$
$$\lambda_2 = \left[P_G\left(F_{GOOD}(T_2) - F_{GOOD}(S_1)\right) + P_B\left(F_{BAD}(T_2) - F_{BAD}(S_1)\right)\right] \cdot \lambda_{COLLECT},$$
$$\lambda_3 = \left[P_G\left(F_{GOOD}(S_1) - F_{GOOD}(S^F)\right) + P_B\left(F_{BAD}(S_1) - F_{BAD}(S^F)\right)\right] \cdot \lambda_{COLLECT}.$$

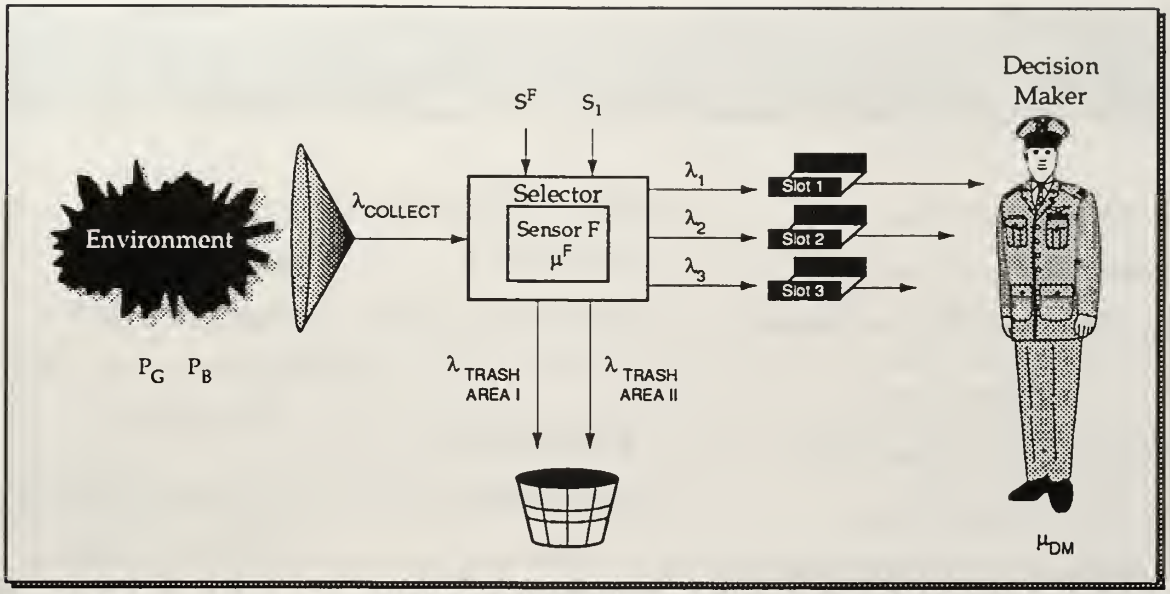

Figure 11. Single Trait Flow Reduction with Prioritized In-Box Model

Table 5 provides the example data for the single trait model. Table 13 adds the additional information required for MOE evaluation. The probability and analysis time MOEs do not change from the single trait model and are located in Table 6 and Table 7 respectively. The waiting time MOEs are presented in Table 14.

EXAMPLE DATA FOR EVALUATION OF MOEs		
Parameter	Equation	Value
Cutoff trait level for In-Box.	S_1	0.77
S_1 set to minimize $W_{DM\ GOOD}$		

Table 13. Priority In-Box Single Trait Flow Reduction Model Example

WAITING TIME MEASURES OF EFFECTIVENESS		
MOE	Equation	Example Result
$W_{\text{DM GOOD}}$ Expected total time in the system for good items reaching the decision maker.	Provided in Appendix A.	0. 466 Hrs. *(Single Trait = 0.55)* *(Base-Case = 0.50)*
$W_{\text{TYPE II ERROR}}$ Expected total time in the system for bad items reaching the decision maker.	Provided in Appendix A.	0.793 Hrs. *(Single Trait = 0.55)* *(Base-Case = 0.50)*
$W_{\text{TRASH BAD}}$ Expected total time in the system for bad items which are trashed.	$\frac{1}{\mu^{F} - \lambda_{\text{COLLECT}}}$	0. 05 Hrs. *(Single Trait = 0.05)* *(Base-Case = 0.0)*
$W_{\text{TYPE I ERROR}}$ Expected total time in the system for good items which are trashed.	$\frac{1}{\mu^{F} - \lambda_{\text{COLLECT}}}$	0.05 Hrs. *((Single Trait = 0.05)* *Base-Case = 0.0)*

Table 14. Priority In-Box Single Trait Flow Reduction Model Waiting Time MOEs

IV. CALIBRATION OF SELECTORS AND SYSTEM SENSITIVITY

A. CALIBRATION OF SELECTOR CUTOFFS

1. Setting the Single Trait Selector to Achieve the Desired Flow Reduction

S^F is the trait level for the single trait selector system above which all items are routed to the decision maker. Recall,

$$\lambda_{DM} = \frac{\rho_{DM}}{\rho_{COLLECT}} \cdot \lambda_{COLLECT},$$

and that the items above S^F are both good and bad with CDFs $F_{GOOD}(t)$ and $F_{BAD}(t)$ respectively. Therefore, the fraction of flow allowed to reach the decision maker must be $\rho_{DM} / \rho_{COLLECT}$.

Select S^F such that,

$$\frac{\rho_{DM}}{\rho_{COLLECT}} = \left(P_G \cdot \left[1 - F_{GOOD}\left(S^F\right)\right] + P_B \cdot \left[1 - F_{BAD}\left(S^F\right)\right]\right).$$

Note that to achieve λ_{DM} only a single S^F setting is available as illustrated in Figure 12. The lower graph plots the resultant flow to the decision maker against the setting S^F. S^F can assume values from T^F_1 (0.1) to T^F_2 (1.0). By setting S^F to T^F_1 the only flow removed from the system is the small fraction of known bad items with trait levels less than T^F_1. Setting S^F to T^F_2 removes all flow from the system except for the known good items which have trait levels above T^F_2. Because the decision maker in the example has the capacity to analyze a total flow of 10 items per hour but also desires to use only 80 percent of his/her total time analyzing items, the desired flow to the decision maker is 8 items per hour. An S^F setting of 0.701 is therefore required. The upper graph in Figure 12 plots the fraction of the items that reach the decision maker that are good, $\rho_{DM\ GOOD}$, against S^F. Note that this value increases by setting S^F to a higher trait level. However, a higher S^F setting results in a corresponding decrease in flow to the decision maker.

This tradeoff must be understood by the decision maker when calibrating the selector and when determining the fraction of available time to be used analyzing information.

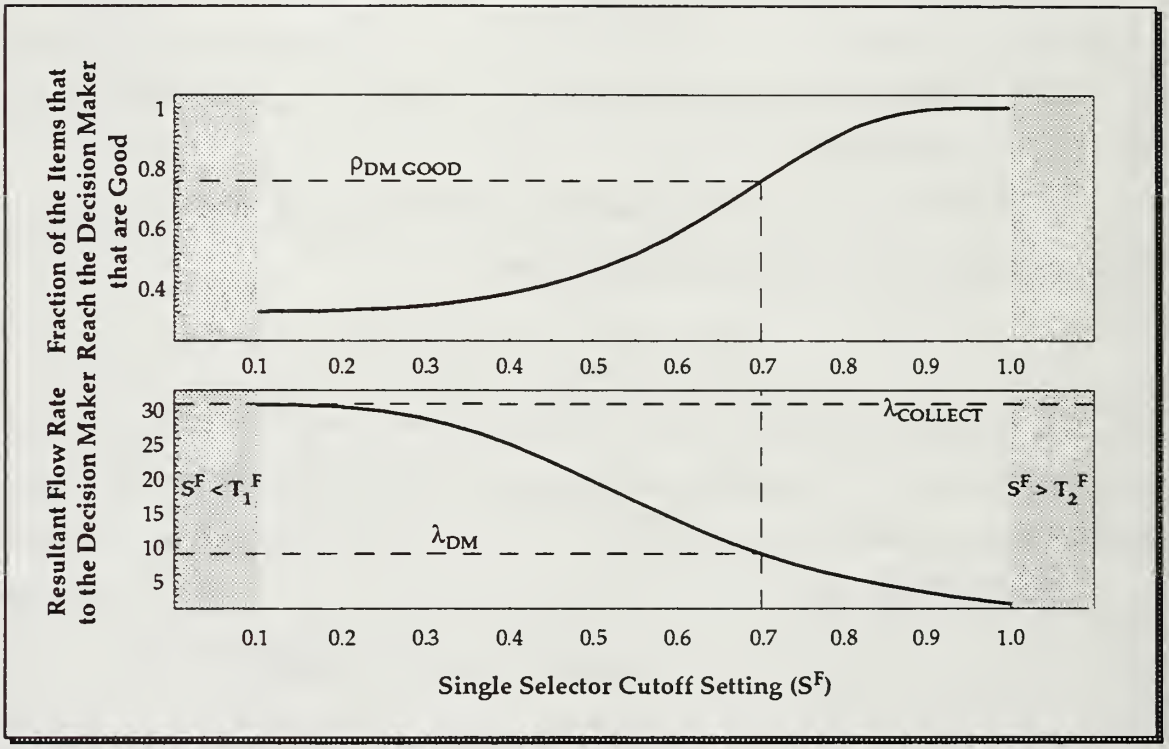

Figure 12. Single Trait Selection Setting to Achieve Flow Reduction and the Corresponding Quality Flow to the Decision Maker

2. Two Trait Selector Settings to Achieve Flow Reduction and Maximize Quality

The dual trait model has two selector settings, S^F and S^G, which must be set to reach the desired reduction in total flow. Unlike the single selector model, however, there are an infinite number of setting combinations which achieve this reduction. If S^F is a setting which allows more flow than the desired λ_{DM} to leave the first selector, S^G can be set to decrease that flow. As demonstrated in the single selector model and in Figure 12, if S^F is set to 0.701 the flow departing the first selector is reduced to λ_{DM}. The second selector, however, will still trash all items with trait level below $T^G{}_1$, Figure 9.

Therefore, the flow reaching the decision maker is less than the desired flow of λ_{DM}. S^F must be set lower than T^F_2 to account for this reduction. Clearly, the setting S^G is uniquely determined once S^F has been set. The upper graph in Figure 13 plots S^G against S^F denoting all valid combinations of the settings which reduce $\lambda_{COLLECT}$ to λ_{DM}. Recalling that the decision maker desires the highest quality of information possible given the available settings, the lower graph in Figure 13 plots the quality achieved, $\rho_{DM\ GOOD}$, versus S^F. A maximum value of 0.805 occurs when S^F equals 0.635. The S^G setting which corresponds to an S^F of 0.635 is 0.3783.

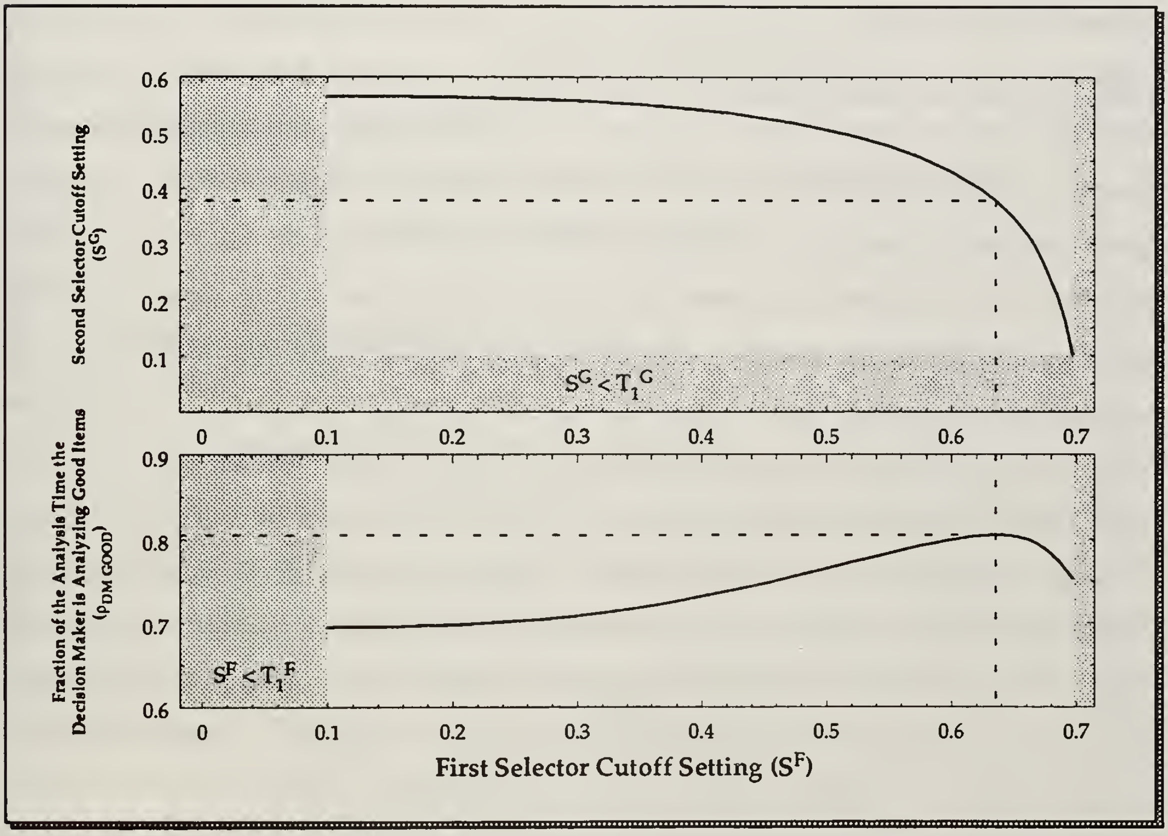

Figure 13. Dual Trait Selector Settings to Achieve Flow Reduction and Maximize Quality

B. CALIBRATION OF A PRIORITY IN-BOX

The priority in-box model uses the single trait selector model as its basis. λ_{DM} is achieved by setting S^F to 0.701. All items leaving the selector with trait levels above T^F_2 are placed in slot 1 of the in-box and have the highest priority, Figure 10. The remaining flow has trait levels between S^F and T^F_2. These items must be placed in either slot 2 or slot 3 as discussed in the model's development. The cutoff setting S_1 is the setting above which items are placed into slot 2 and below which items are placed into slot 3. The role of the prioritized in-box is to reduce the expected waiting time in the system for good items by reducing the expected waiting time in the decision maker's in-box. Therefore, S_1 is selected to provide the lowest expected waiting time for good items. Figure 14 plots the expected waiting time for good and bad items versus S_1. The solid line is the waiting time for good items and has a minimum value of 0.466 hours when S_1 is set to 0.77. Note that this is well below the result of 0.55 hours for the single selector model without a prioritized in-box. The expected waiting times for bad items are plotted with a dashed line. Note that for all settings of S_1 the bad items take longer to pass through the system than the good items. When S_1 is set to a level close to the lower bound of slot 3 or the lower bound of slot 1 the system behaves similarly to a two slot in-box. The separation of the curves at the boundaries results from the priority of slot 1 and that all items in slot 1 are good. At optimum this time separation allows good items to reach the decision maker 40 percent faster than bad items. Therefore there is never a disadvantage to prioritizing the flow to the decision maker.

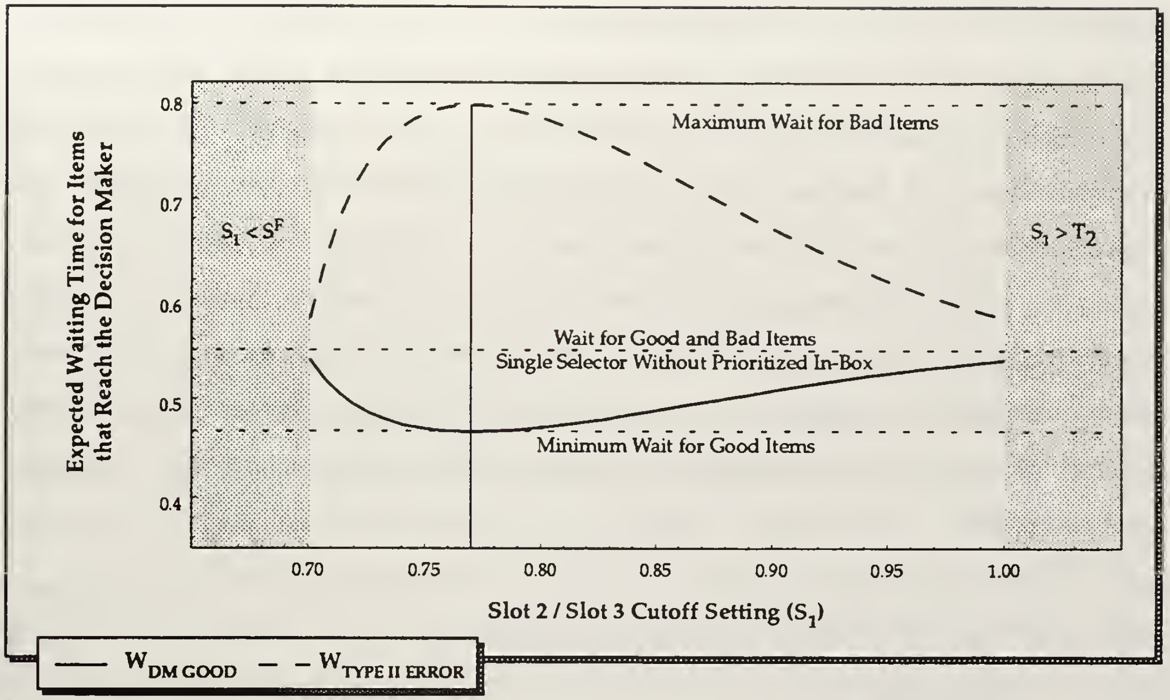

Figure 14. Single Selector with Prioritized In-Box Model
Waiting Times versus S_1

C. SYSTEM SENSITIVITY ANALYSIS

1. Two Trait Selector Sensitivity to $\lambda_{COLLECT}$

It is important to understand the implications of not recalibrating a flow reduction system when changes occur external to the system. Figure 15 displays 8 quality curves plotted against S^F. The dark curve provides the decision maker with 8 items per hour as discussed in model development and calibration. The lighter curves correspond to amounts from 10 to 3 items per hour as indicated. 10 items per hour is the maximum rate of flow a decision maker can analyze with the given service rate of $\mu_{DM} = 10$. A higher rate causes information overload. Also at $\lambda_{DM} = 10$ the decision maker is spending all of his/her available time analyzing items ($\rho_{DM} = 1.0$) and no time making decisions. For the example system, S^F cannot be set to a level which allows for 2 or fewer items per hour. Note that, if the model setting for S^F (0.635 as indicated by a vertical line) is retained and the arrival rate to the

decision maker declines due to a decrease in $\lambda_{COLLECT}$, there is an increase in the quality of arriving items. This may occur when the rate at which items are collected decreases but the environment from which the items are obtained does not change. Also note that though the quality is higher, the setting does not provide the highest quality possible for that rate of arrivals. In this case S^F should be reset to achieve optimum quality. A similar decline in quality occurs when λ_{DM} increases through an increase in $\lambda_{COLLECT}$ and a new optimum quality is reached by decreasing S^F. Therefore, the actual collection rate must be monitored and the system recalibrated to maintain optimum quality levels.

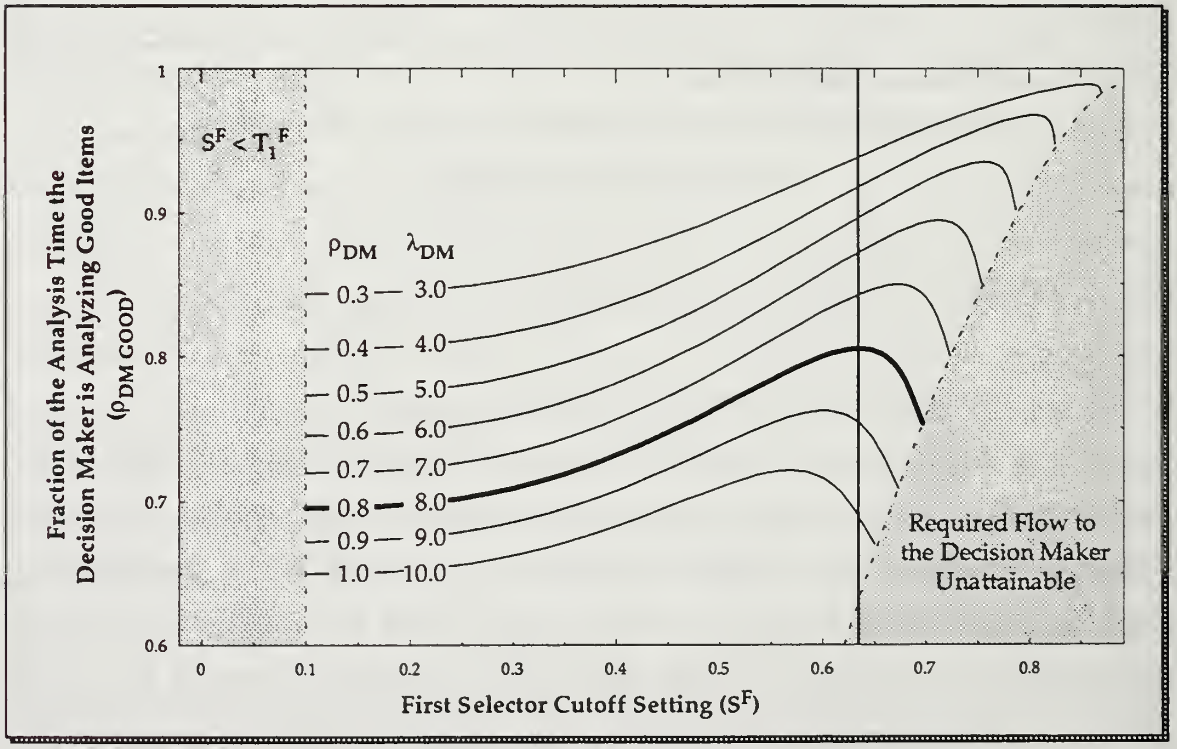

Figure 15. Recalibration Requirement Due to Changes in $\lambda_{COLLECT}$

2. Model Sensitivity to Changes in the Probability a Collected Item is Good

Changes in the information environment as well as changes in the collection system may induce a change in the probability that a collected item

is good, P_G. This change in P_G is directly reflected in the quality of information reaching the decision maker as measured by $\rho_{DM\ GOOD}$. Figure 16 demonstrates this effect. The figure plots $\rho_{DM\ GOOD}$ against P_G for both the single and dual trait selector systems. Note that over the entire range of P_G the dual trait system provides higher quality information to the decision maker and both systems are better than the non-discriminating base-case. If a system produced lower quality than base-case (the light gray area) it would not be used. Recall that the example data calibrated the system based upon an estimated P_G value of 0.30. Indicator lines are plotted at $P_G = 0.2$ and $P_G = 0.4$ to indicate a region of uncertainty in the actual value of P_G. The resultant $\rho_{DM\ GOOD}$ value levels are also plotted. For the single trait selector system, this uncertainty results in $\rho_{DM\ GOOD}$ values ranging from 0.5 to 0.725. The same uncertainty imposed upon the dual trait model results in $\rho_{DM\ GOOD}$ values from about 0.7 to 0.86. Therefore the dual trait system is much less sensitive to fluctuations in P_G in the estimated region.

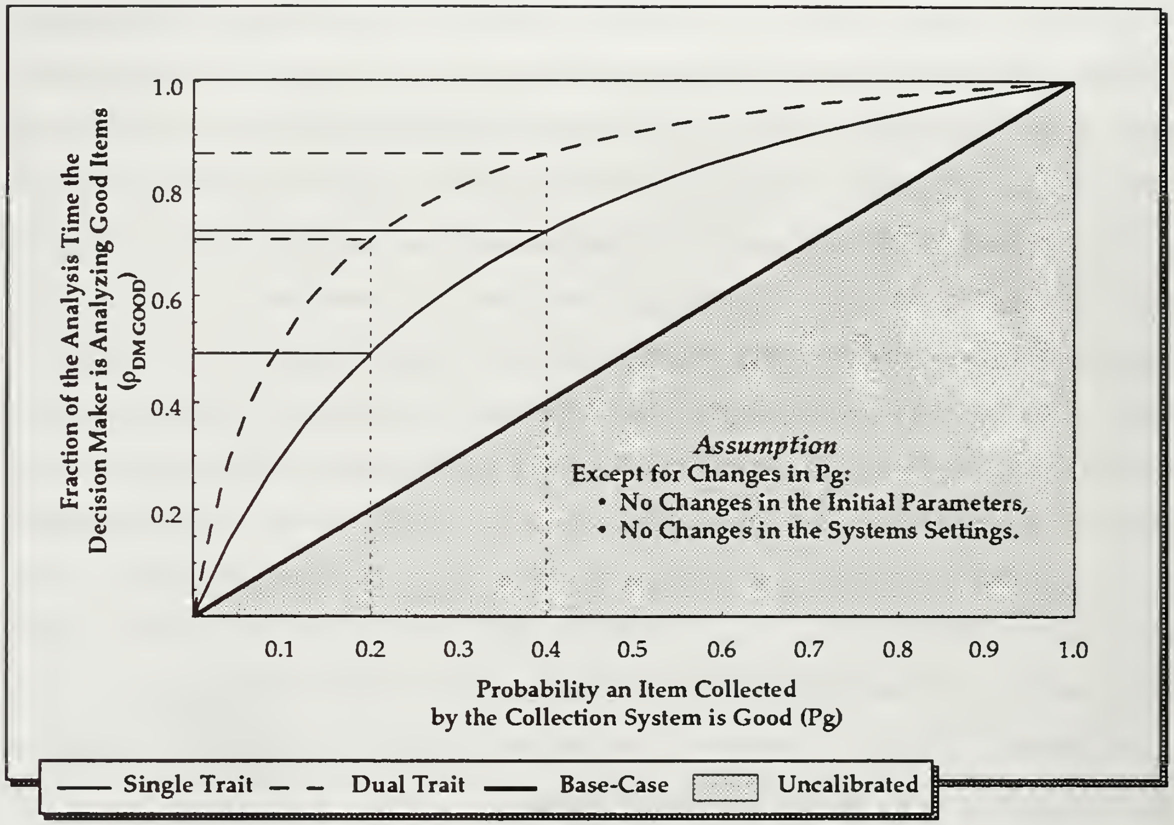

Figure 16. Impact of Changes in P_G on $\rho_{DM\ GOOD}$ for Single and Dual Trait Systems

The same curves as in Figure 16 are shown in Figure 17 to demonstrate that they can be used to evaluate the impact of adding an additional selector to the system or adjusting the collection system or the information environment to achieve higher levels of information quality. In this case we assume the decision maker has the ability to adjust the collection system and/or in some fashion change the environment to increase P_G. It is assumed that these additions or changes result in a cost to the decision maker which can be estimated. Using the example data and a single trait system, a $\rho_{DM\ GOOD}$ value of about 0.74 is achieved at a P_G of 0.3 (see Figure 17). Suppose the decision maker desires this quality to be increased to 0.8. The addition of a second selection achieves this value directly. If a selector is not added to the system, the decision maker must change the collection system and/or the information environment such that the new P_G value is approximately 0.5.

The costs associated with each change can be compared and the most cost effective option chosen.

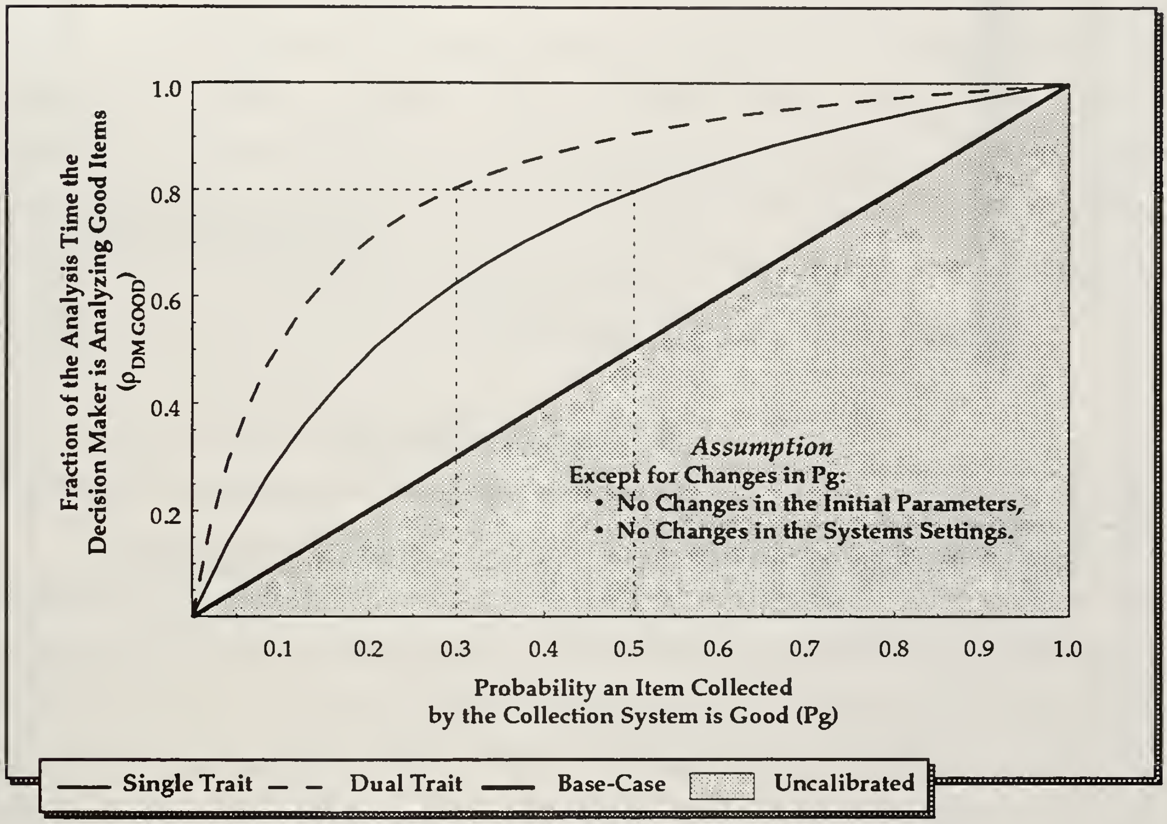

Figure 17. Adding a Second Selector or Adjusting P_G

The priority in-box is sensitive to changes in P_G as illustrated in Figure 18. As in Figure 14, the expected time in the system for both good and bad items is plotted against S_1. The four sets of curves show the expected delay if the system receives information from the environment at P_G levels of 0.1, 0.3, 0.5 and 0.7 respectively. The system is assumed to have been recalibrated following the change in P_G. Note that as P_G decreases the delay in the system decreases for both good and bad items. The delay for good items approaches that of a non-priority in-box as P_G approaches 1.0. Therefore there is no disadvantage across P_G level to adding a priority in-box to a flow reduction system but a great advantage if P_G is small.

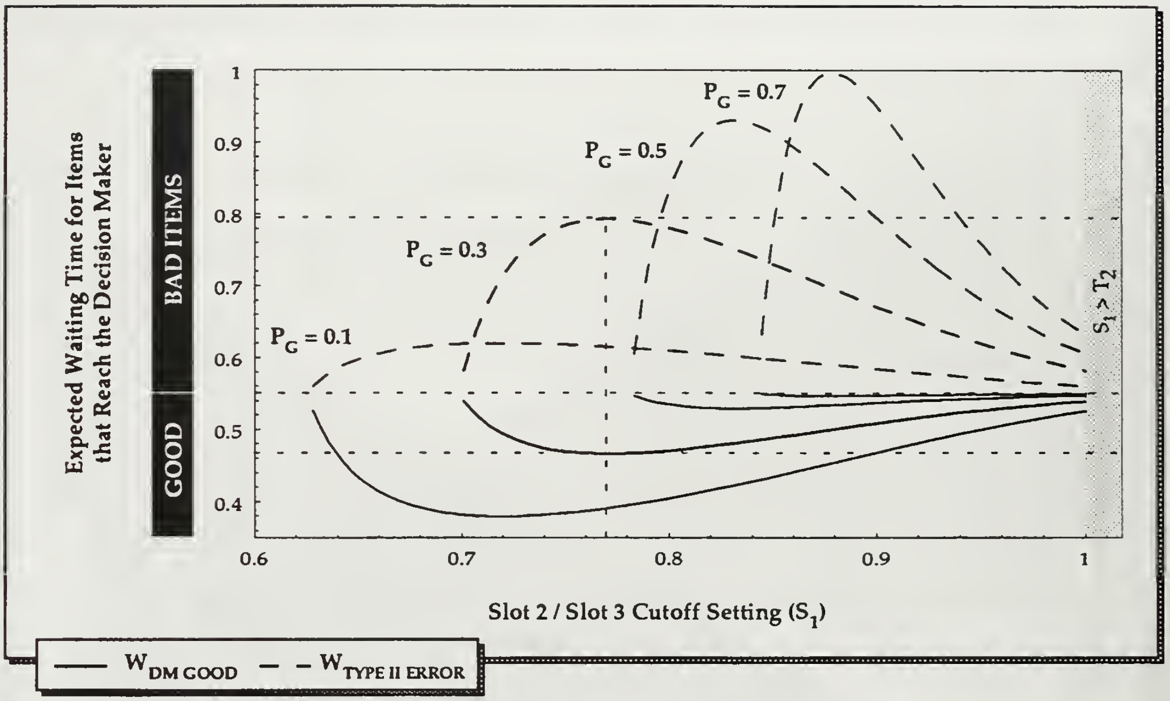

Figure 18. Sensitivity of a Priority In-Box System to Changes in P_G

V. CONCLUSIONS

This thesis develops several analytical models for the evaluation of an information flow reduction system based upon probabilistic inputs. Sensitivity analyses demonstrate the requirement to monitor the estimated parameters of the system in order to maintain flow to the decision maker which is both acceptable and optimal. The study closes with some comments on the significance of the models and suggestion for further research.

A. MODEL INSIGHTS AND USEFULNESS

The four models presented provide an insight into the capabilities and limitations of probabilistic based flow reduction systems. Such systems can provide high quality information to the decision maker at the desired flow rate if traits can be found which provide separation and/or skewness between the value probability curves. Such systems require historical records to evaluate the information environment and are subject to fluctuations in the parameters of that environment. This implies two important concepts:

- an understanding of the enemy's methods and tactics is fundamental to operating these systems and,
- the enemy can take advantage of friendly use of these systems by changing their operating parameters or hiding truly vital information in low probability regions.

Selector-based systems can increase flow quality with a trade off in delay to the operator. The models utilize machine-based sensors to sample the trait level of an item. Therefore, in most cases, the delay is not significant when compared with the human decision maker. This may not always be the case and the decision maker must take into account the service rates of all "devices" in the reduction system, some of which may be human as well.

The priority in-box demonstrates the ability of the system to decrease the delay of good items reaching the decision maker. This decrease occurs across every calibration setting. Therefore there is no disadvantage to implementing a prioritizing system. In the models the prioritizing system was placed just before the decision maker but this does not need to be the case. A prioritized in-box can be used following every selector in the reduction system and increase the speed of good items across the entire network.

B. RECOMMENDATIONS FOR FURTHER STUDY

Further study is warranted in two distinct areas. The first area deals with the operation of the system to maintain optimal conditions. The second focuses on managers of these systems and anticipated developments.

In the development of this thesis an assumption was made which stated that the parameters of the information environment did not change and the system was allowed to attain steady state. However, analysis revealed that these systems are susceptible to fluctuations in those parameters and need recalibration if the parameters change. A procedure should be developed to monitor the changes in the information environment to include $\lambda_{COLLECT}$, P_G, $f_{GOOD}(t)$ and $f_{BAD}(t)$ (for each monitored trait).

Managers of such systems must anticipate changes in the requirements associated with the system. Procedures should be developed to answer management questions such as:

- If an investment is to be made, what portion of the system should be improved/replaced?
- Is the output quality of the system as high as the system is capable of achieving? How do we know?
- What is the limiting device in the system?
- Can this system achieve my desires?
- Is the system doing what I expect? How do we know?

- Do we have the best quality/quantity mix?
- If we have to remove a device, which should it be?
- Are the selectors in the optimum order?

Tactical questions involve the exploitation of such systems. How can the enemy exploit our use of these systems? How can we exploit the enemy's use of them? What information about these systems is vital to suppress from the enemy and thus what information is important for us to obtain about theirs?

APPENDIX. DERIVATIONS OF MEASURE OF EFFECTIVENESS

1. NONDISCRIMINATING SELECTION PROCESS

A. $P_{DM\ GOOD}$

$P_{DM\ GOOD}$ is the probability an item is good and it reaches the decision maker. The probability any item reaches the decision maker is

$$\alpha = \frac{\lambda_{DM}}{\lambda_{COLLECT}} = \frac{\rho_{DM}}{\rho_{COLLECT}}.$$

The probability an item is good given that it is selected is

$$P\{Value = Good \mid Selected\} = \frac{P\{Value = Good, Selected\}}{P\{Selected\}},$$

and by independence,

$$\begin{aligned} P\{Value = Good \mid Selected\} &= \frac{P\{Value = Good\} \cdot P\{Selected\}}{P\{Selected\}} \\ &= P\{Value = Good\} \\ &= P_G. \end{aligned}$$

Therefore, $P_{DM\ GOOD}$ equals

$$\begin{aligned} P_{DM\ GOOD} &= P\{Value = Good \mid Selected\} \cdot P\{Selected\} \\ &= P_G \cdot \left(\frac{\rho_{DM}}{\rho_{COLLECT}} \right). \end{aligned}$$

B. $P_{\text{TYPE I ERROR}}$

$P_{\text{TYPE I ERROR}}$ is the probability that an item is good and it is trashed. The probability any item is trashed is

$$1-\alpha=\left(1-\frac{\rho_{\text{DM}}}{\rho_{\text{COLLECT}}}\right).$$

The probability an item is good given that it is trashed is

$$\mathrm{P}\{\text{Value} = \text{Good} \mid \text{Trashed}\} = \frac{\mathrm{P}\{\text{Value} = \text{Good, Trashed}\}}{\mathrm{P}\{\text{Trashed}\}},$$

and by independence,

$$\begin{aligned}\mathrm{P}\{\text{Value} = \text{Good} \mid \text{Trashed}\} &= \frac{\mathrm{P}\{\text{Value} = \text{Good}\}\cdot \mathrm{P}\{\text{Trashed}\}}{\mathrm{P}\{\text{Trashed}\}} \\ &= \mathrm{P}\{\text{Value} = \text{Good}\} \\ &= \mathrm{P}_{\mathrm{G}} \quad .\end{aligned}$$

Therefore, $P_{\text{TYPE I ERROR}}$ equals

$$\begin{aligned}\mathrm{P}_{\text{TYPE I ERROR}} &= \mathrm{P}\{\text{Value} = \text{Good} \mid \text{Trashed}\}\cdot \mathrm{P}\{\text{Trashed}\} \\ &= \mathrm{P}_{\mathrm{G}}\cdot\left(1-\frac{\rho_{DM}}{\rho_{\text{COLLECT}}}\right).\end{aligned}$$

C. $P_{\text{TRASH BAD}}$

$P_{\text{TRASH BAD}}$ is the probability that an item is good and it is trashed. Therefore, $P_{\text{TRASH BAD}}$ equals

$$\begin{aligned} P_{\text{TRASH BAD}} &= P\{\text{Value} = \text{Bad} \mid \text{Trashed}\} \cdot P\{\text{Trashed}\} \\ &= P_B \cdot \left(1 - \frac{\rho_{DM}}{\rho_{\text{COLLECT}}}\right). \end{aligned}$$

D. $P_{\text{TYPE II ERROR}}$

$P_{\text{TYPE II ERROR}}$ is the probability that an item is bad and it is selected. Therefore, $P_{\text{TYPE II ERROR}}$ equals

$$\begin{aligned} P_{\text{TYPE II ERROR}} &= P\{\text{Value} = \text{Bad} \mid \text{Selected}\} \cdot P\{\text{Selected}\} \\ &= P_B \cdot \left(\frac{\rho_{DM}}{\rho_{\text{COLLECT}}}\right). \end{aligned}$$

E. $\rho_{\text{DM GOOD}}$

$\rho_{\text{DM GOOD}}$ is the fraction of the decision maker's time spent on good items. This is the fraction of items that reach the decision maker that are good. The fraction of all item that are good and reach the decision maker is defined to be $P_{\text{DM GOOD}}$. The fraction of all items that are bad and reach the decision maker is defined to be $P_{\text{TYPE II ERROR}}$. Therefore, the fraction of those items that reach the decision maker that are good is

$$\frac{P_{\text{DM GOOD}}}{P_{\text{DM GOOD}} + P_{\text{TYPE II ERROR}}}.$$

F. $\rho_{\text{TYPE II ERROR}}$

The same derivation is used to determine $\rho_{\text{TYPE II ERROR}}$, which becomes,

$$\frac{P_{\text{TYPE II ERROR}}}{P_{\text{DM GOOD}} + P_{\text{TYPE II ERROR}}}.$$

G. $W_{\text{DM GOOD}}$ and $W_{\text{TYPE II ERROR}}$

Items that reach the decision maker have two opportunities to be delayed. The first delay is from the selection process and the second delay is waiting for and then being analyzed by the decision maker. The selector delay is D and is very small when compared to μ_{DM}. Therefore the only delay is that seen at the decision maker. This is the delay for a single channel M/M/1 queueing system and the result is

$$\frac{1}{\mu_{DM} - \lambda_{DM}}.$$

H. $W_{\text{TRASH BAD}}$ and $W_{\text{TYPE I ERROR}}$

Items that do not reach the decision maker have one opportunity to be delayed; the selection process. The selector delay is D and is very small when compared to μ_{DM} and is considered to be zero.

2. SINGLE TRAIT ANALYSIS WITHOUT PRIORITY IN-BOX

A. $P_{\text{DM GOOD}}$

$P_{\text{DM GOOD}}$ is the probability an item is good and it reaches the decision maker. Good items reach the decision maker only if they possess trait levels greater than S^F. Therefore, $P_{\text{DM GOOD}}$ equals

$$\begin{aligned}
P_{\text{DM GOOD}} &= P\{\text{Selected} \mid \text{Good}\} \cdot P\{\text{Good}\} \\
&= P_G \cdot \int_{S^F}^{T_3} f_{GOOD}(s)ds \\
&= P_G \cdot \left(1 - F_{GOOD}\left(S^F\right)\right).
\end{aligned}$$

B. $P_{\text{TYPE I ERROR}}$

$P_{\text{TYPE I ERROR}}$ is the probability that an item is good and it is trashed. All items with trait values below S^F are trashed. Therefore, $P_{\text{TYPE I ERROR}}$ equals

$$\begin{aligned}
P_{\text{TYPE I ERROR}} &= P\{\text{Trashed} \mid \text{Good}\} \cdot P\{\text{Good}\} \\
&= P_G \cdot \int_{T_1}^{S^F} f_{GOOD}(s)ds \\
&= P_G \cdot F_{GOOD}\left(S^F\right).
\end{aligned}$$

C. $P_{\text{TRASH BAD}}$

$P_{\text{TRASH BAD}}$ is the probability that an item is good and it is trashed. All items with trait values below S^F are trashed. Therefore, $P_{\text{TRASH BAD}}$ equals

$$\begin{aligned}
P_{\text{TRASH BAD}} &= P\{\text{Trashed} \mid \text{Bad}\} \cdot P\{\text{Bad}\} \\
&= P_B \cdot \int_{T_0}^{S^F} f_{BAD}(s)ds \\
&= P_B \cdot F_{BAD}\left(S^F\right).
\end{aligned}$$

D. $P_{\text{TYPE II ERROR}}$

$P_{\text{TYPE II ERROR}}$ is the probability that an item is bad and it is selected. Bad items reach the decision maker only if they possess trait levels greater than S^F. Therefore, $P_{\text{TYPE II ERROR}}$ equals

$$
\begin{aligned}
P_{\text{DM GOOD}} &= P\{\text{Selected} \mid \text{Bad}\} \cdot P\{\text{Bad}\} \\
&= P_B \cdot \int_{S^F}^{T_2} f_{BAD}(s)\,ds \\
&= P_B \cdot \left(1 - F_{BAD}\left(S^F\right)\right).
\end{aligned}
$$

E. $\rho_{\text{DM GOOD}}$

$\rho_{\text{DM GOOD}}$ is the fraction of the decision maker's analysis time spent on good items. This is the fraction of items that reach the decision maker that are good. The fraction of all items that are good and reach the decision maker is defined to be $P_{\text{DM GOOD}}$. The fraction of all items that are bad and reach the decision maker is defined to be $P_{\text{TYPE II ERROR}}$. Therefore, the fraction of items which reach the decision maker that are good is

$$
\frac{P_{\text{DM GOOD}}}{P_{\text{DM GOOD}} + P_{\text{TYPE II ERROR}}}.
$$

F. $\rho_{\text{TYPE II ERROR}}$

The same derivation is used to determine $\rho_{\text{TYPE II ERROR}}$, which becomes,

$$
\frac{P_{\text{TYPE II ERROR}}}{P_{\text{DM GOOD}} + P_{\text{TYPE II ERROR}}}.
$$

G. $W_{\text{DM GOOD}}$ and $W_{\text{TYPE II ERROR}}$

Items that reach the decision maker have two opportunities to be delayed. The first delay is from the selection process and the second delay is waiting for and then being analyzed by the decision maker. The selector delay is exponential with rate μ^F. The decision maker delay is also exponential with rate μ_{DM}. Both delays form single channel M/M/1 queueing systems and the result is

$$\frac{1}{\mu^F - \lambda_{\text{COLLECT}}} + \frac{1}{\mu_{\text{DM}} - \lambda_{\text{DM}}}.$$

H. $W_{\text{TRASH BAD}}$ and $W_{\text{TYPE I ERROR}}$

Items that do not reach the decision maker have one opportunity to be delayed; the selection process. The selector delay is exponential with rate μ^F. This delay forms a single channel M/M/1 queueing system and the result is

$$\frac{1}{\mu_{\text{DM}} - \lambda_{\text{DM}}}.$$

3. TWO TRAIT ANALYSIS

A. $P_{DM\ GOOD}$

$P_{DM\ GOOD}$ is the probability an item is good and it reaches the decision maker. In the two trait selector system, good items can reach the decision maker directly from either selector. Good items reach the decision maker directly from the first selector only if they possess trait levels greater than T_2^F. This value is

$$= P_G \cdot \left(\int_{T_2^F}^{T_3^F} f_{GOOD}(s)ds \right) = P_G \cdot \left(1 - F_{GOOD}(T_2^F)\right).$$

Good items reach the decision maker from the second selector only if they are selected by the first selector but not sent directly to the decision maker and then are selected by the second selector. Define P_{FTOG} as the probability an item is sent from the first selector to the second selector. Then

$$P_{F\,TO\,G} = \left(P_G \cdot \left[F_{GOOD}(T_2^F) - F_{GOOD}(S^F)\right] + P_B \cdot \left[F_{BAD}(T_2^F) - F_{BAD}(S^F)\right]\right).$$

Using this notation we calculate the probability an arrival to the second selector is good or bad. Define these probabilities as $P_{GOOD\ |\ F\,TO\,G}$ and $P_{BAD\ |\ F\,TO\,G}$ where,

$$P_{GOOD\ |\ F\,TO\,G} = \frac{P_{SELECTED\ GOOD\ F\ AREA\ II}}{P_{F\,TO\,G}}$$

$$= \frac{P_G \cdot \left(\int_{S^F}^{T_2^F} f_{GOOD}(s) \cdot ds \right)}{\left(P_G \cdot \left[F_{GOOD}(T_2^F) - F_{GOOD}(S^F)\right] + P_B \cdot \left[F_{BAD}(T_2^F) - F_{BAD}(S^F)\right]\right)}$$

$$= \frac{P_G \cdot \left[F_{GOOD}(T_2^F) - F_{GOOD}(S^F)\right]}{\left(P_G \cdot \left[F_{GOOD}(T_2^F) - F_{GOOD}(S^F)\right] + P_B \cdot \left[F_{BAD}(T_2^F) - F_{BAD}(S^F)\right]\right)}$$

and

$$P_{BAD\,|\,F\,TO\,G} = \frac{P_{SELECTED\ BAD\ F\ AREA\ II}}{P_{F\,TO\,G}}$$

$$= \frac{P_B \cdot \left(\int_{S^F}^{T_2^F} f_{BAD}(s) \cdot ds \right)}{\left(P_G \cdot \left[F_{GOOD}(T_2^F) - F_{GOOD}(S^F)\right] + P_B \cdot \left[F_{BAD}(T_2^F) - F_{BAD}(S^F)\right]\right)}$$

$$= \frac{P_B \cdot \left[F_{BAD}(T_2^F) - F_{BAD}(S^F)\right]}{\left(P_G \cdot \left[F_{GOOD}(T_2^F) - F_{GOOD}(S^F)\right] + P_B \cdot \left[F_{BAD}(T_2^F) - F_{BAD}(S^F)\right]\right)}.$$

For all arrivals to the second selector it acts like the single selector in the single trait case. Therefore, $P_{DM\ GOOD}$ equals

$$\left[P_G \cdot \left(1 - F_{GOOD}(T_2^F)\right)\right] + \left[P_{F\,TO\,G} \cdot P_{GOOD\,|\,F\,TO\,G} \cdot \left(1 - G_{GOOD}(S^G)\right)\right].$$

B. $P_{TYPE\ I\ ERROR}$

$P_{TYPE\ I\ ERROR}$ is the probability that an item is good and it is trashed. In the first selector, all items with trait values below S^F are trashed. This probability was calculated in the single trait case to be

$$P_G \cdot F_{GOOD}(S^F).$$

For the second selector the items must arrive to the selector, be good items and be trashed. Items with trait values less than S^G are trashed, therefore the total value is

$$\left[P_G \cdot F_{GOOD}(S^F)\right] + \left[P_{F\,TO\,G} \cdot P_{GOOD\,|\,F\,TO\,G} \cdot G_{GOOD}(S^G)\right].$$

C. $P_{TRASH\ BAD}$

$P_{TRASH\ BAD}$ is the probability that an item is good and it is trashed. In the first selector all items with trait values below S^F are trashed and in the second selector all items with trait levels below S^G are trashed. Therefore, $P_{TRASH\ BAD}$ equals

$$\left[P_B \cdot F_{BAD}\left(S^F\right)\right] + \left[P_{F\,TO\,G} \cdot P_{BAD\,|\,F\,TO\,G} \cdot G_{BAD}\left(S^G\right)\right].$$

D. $P_{TYPE\ II\ ERROR}$

$P_{TYPE\ II\ ERROR}$ is the probability that an item is bad and it is selected. Bad items reach the decision maker only if they possess trait levels greater than S^F in the first selector, are sent to the second selector, and have trait levels greater than S^G in the second selector. Therefore, $P_{TYPE\ II\ ERROR}$ equals

$$\left[P_{F\,TO\,G} \cdot P_{BAD\,|\,F\,TO\,G} \cdot \left(1 - G_{BAD}\left(S^G\right)\right)\right].$$

E. $\rho_{DM\ GOOD}$

$\rho_{DM\ GOOD}$ is the fraction of the decision maker's time spent on good items. This is the fraction of items that reach the decision maker that are good. The fraction of all items that are good and reach the decision maker is defined to be $P_{DM\ GOOD}$. The fraction of all items that are bad and reach the decision maker is defined to be $P_{TYPE\ II\ ERROR}$. Therefore, the fraction of items which reach the decision maker that are good is

$$\frac{P_{DM\ GOOD}}{P_{DM\ GOOD} + P_{TYPE\ II\ ERROR}}.$$

F. $\rho_{\text{TYPE II ERROR}}$

The same derivation is used to determine $\rho_{\text{TYPE II ERROR}}$, which becomes,

$$\frac{P_{\text{TYPE II ERROR}}}{P_{\text{DM GOOD}} + P_{\text{TYPE II ERROR}}}.$$

G. $W_{\text{DM GOOD}}$

Items that reach the decision maker have three opportunities to be delayed. The first delay is from the first selection process and all items must pass this process. The second delay is for the second selection process. Only those items that have trait levels above S^F but below T_2^F are sent to the second selector. The remaining items, if they are good, are sent directly to the decision maker and do not see this delay. Once reaching the decision maker, all good items must wait for and then be analyzed by the decision maker. The first selector service time is distributed exponential with rate μ^F. The second selector service time is distributed exponential with rate μ^G. Only a fraction of items that are collected reach the second selector. This fraction is P_{FTOG} as previously calculated. Define

$$\lambda_G^F = P_{FTOG} \cdot \lambda_{COLLECT}.$$

The decision maker's service time is distributed exponential with rate μ_{DM}. The flow reaching the decision maker has been reduced to λ_{DM}. All devices form a single channel M/M/1 queueing systems. The delay added by the second selector must be weighted to include only the good items that pass through it. This weighting is

$$\left(\frac{P_{FTOG} \cdot P_{GOOD\,|\,FTOG} \cdot \left(1 - G_{GOOD}(S^G)\right)}{P_G \cdot \left(1 - F_{GOOD}(T_2^F)\right) + \left[P_{FTOG} \cdot P_{GOOD\,|\,FTOG} \cdot \left(1 - G_{GOOD}(S^G)\right)\right]}\right).$$

Therefore $W_{DM\ GOOD}$ is

$$\frac{1}{\mu^{F}-\lambda_{COLLECT}}+\left(\frac{P_{F\ TO\ G}\cdot P_{GOOD\ |\ F\ TO\ G}\cdot\left(1-G_{GOOD}\left(S^{G}\right)\right)}{P_{G}\cdot\left(1-F_{GOOD}\left(T_{2}^{F}\right)\right)+\left[P_{F\ TO\ G}\cdot P_{GOOD\ |\ F\ TO\ G}\cdot\left(1-G_{GOOD}\left(S^{G}\right)\right)\right]}\right)\cdot\left(\frac{1}{\mu^{G}-\lambda_{G}^{F}}\right)+\left(\frac{1}{\mu_{DM}-\lambda_{DM}}\right).$$

H. $W_{TYPE\ II\ ERROR}$

All bad items that reach the decision maker must pass both selectors and wait to be analyzed by the decision maker. Therefore their waiting time is

$$\frac{1}{\mu^{F}-\lambda_{COLLECT}}+\left(\frac{1}{\mu^{G}-\lambda_{G}^{F}}\right)+\left(\frac{1}{\mu_{DM}-\lambda_{DM}}\right).$$

I. $W_{TRASH\ BAD}$

Bad items are trashed by both the first and second selectors. Only a fraction of the bad trashed items are trashed in the second selector. Therefore, the waiting time associated with the second selector is weighted accordingly with the equation

$$\left(\frac{P_{F\ TO\ G}\cdot P_{BAD\ |\ F\ TO\ G}\cdot G_{BAD}\left(S^{G}\right)}{\left[P_{F\ TO\ G}\cdot P_{BAD\ |\ F\ TO\ G}\cdot G_{BAD}\left(S^{G}\right)\right]+P_{B}\cdot F_{BAD}\left(S^{F}\right)}\right).$$

Therefore, $W_{TRASH\ BAD}$ is

$$\left(\frac{1}{\mu^{F}-\lambda_{COLLECT}}\right)+\left(\frac{P_{F\ TO\ G}\cdot P_{BAD\ |\ F\ TO\ G}\cdot G_{BAD}\left(S^{G}\right)}{\left[P_{F\ TO\ G}\cdot P_{BAD\ |\ F\ TO\ G}\cdot G_{BAD}\left(S^{G}\right)\right]+P_{B}\cdot F_{BAD}\left(S^{F}\right)}\right)\cdot\left(\frac{1}{\mu^{G}-\lambda_{G}^{F}}\right).$$

J. $W_{TYPE\ I\ ERROR}$

Good items are trashed by both the first and second selector. The fraction of the waiting time associated with the second selector is weighted to account only for those items that reach the second selector. This weight is

$$\left(\frac{P_{FTOG} \cdot P_{GOOD\,|\,FTOG} \cdot G_{GOOD}\left(S^{G}\right)}{\left[P_{FTOG} \cdot P_{GOOD\,|\,FTOG} \cdot G_{GOOD}\left(S^{G}\right)\right] + P_{G} \cdot F_{GOOD}\left(S^{F}\right)}\right).$$

Therefore, $W_{TYPE\ I\ ERROR}$ is

$$\left(\frac{1}{\mu^{F} - \lambda_{COLLECT}}\right) + \left(\frac{P_{FTOG} \cdot P_{GOOD\,|\,FTOG} \cdot G_{GOOD}\left(S^{G}\right)}{\left[P_{FTOG} \cdot P_{GOOD\,|\,FTOG} \cdot G_{GOOD}\left(S^{G}\right)\right] + P_{G} \cdot F_{GOOD}\left(S^{F}\right)}\right) \cdot \left(\frac{1}{\mu^{G} - \lambda_{G}^{F}}\right).$$

4. SINGLE TRAIT ANALYSIS WITH PRIORITY IN-BOX

A. Non-Waiting Time Measures of Effectiveness

All non-waiting time MOEs are quality oriented and are accounted for by the selection process. These MOEs do not change from the single trait selection derivations in Section 2 of this Appendix.

B. $W_{\text{DM GOOD}}$

Good items that reach the decision maker wait at the first selector and the decision maker. The wait for the first selector is

$$\frac{1}{\mu_{\text{DM}} - \lambda_{\text{DM}}}$$

Good items that reach the decision maker are placed in any of the three slots in the priority in-box. The wait in any of these slots is found with the equation, [Ref. 4],

$$W_q^{(i)} = \frac{\sum_{k=1}^{r} \left(\frac{\lambda_k}{\mu_k^2} \right)}{(1-\sigma_{i-1})(1-\sigma_i)}$$

where

r	≡	total number of slots (3),
λ_k	≡	the arrival rate to slot k ε {1, 2, 3},
μ_k	≡	the service rate for items of the type placed in slot k,
	=	μ_{DM} for all slots,
σ_i	≡	$\sum_{j=1}^{i} \frac{\lambda_j}{\mu_j}, \quad (\sigma_0 = 0).$

There is an additional wait for service completion for all items of $1/\mu_{\text{DM}}$. Therefore, for each slot we determine the arrival rate and the probability that

a good item that arrives to the in-box enters that slot. These probabilities are used to weight the expected value of waiting time. Define the slots in order of priority as {1, 2, 3} and their corresponding arrival rates and slot entrance probabilities to be $\{\lambda_1, \lambda_2, \lambda_3\}$ and $\{P_{1GOOD}, P_{2GOOD}, P_{3GOOD}\}$ respectively. Recall that all items entering slot 1 are good and have trait values above T_2^F. Also, the items entering slots 2 and 3 are both good and bad with trait values between S^F and T_2^F and a trait level cutoff between slot 2 and 3 of S_1. Therefore,

$$\lambda_1 = P_G \cdot \left(1 - F_{GOOD}\left(T_2^F\right)\right) \cdot \lambda_{COLLECT},$$

$$\lambda_2 = \left(\left[P_G \cdot \left(F_{GOOD}\left(T_2^F\right) - F_{GOOD}\left(S_1\right)\right)\right] + \left[P_B \cdot \left(1 - F_{BAD}\left(S_1\right)\right)\right]\right) \cdot \lambda_{COLLECT},$$

$$\lambda_3 = \left(\left[P_G \cdot \left(F_{GOOD}\left(S_1\right) - F_{GOOD}\left(S^F\right)\right)\right] + \left[P_B \cdot \left(F_{BAD}\left(S_1\right) - F_{BAD}\left(S^F\right)\right)\right]\right) \cdot \lambda_{COLLECT},$$

$$P_{1GOOD} = \frac{\left[P_G \cdot \left(1 - F_{GOOD}\left(T_2^F\right)\right)\right]}{\left[P_G \cdot \left(1 - F_{GOOD}\left(S^F\right)\right)\right]},$$

$$P_{2GOOD} = \frac{\left[P_G \cdot \left(F_{GOOD}\left(T_2^F\right) - F_{GOOD}\left(S_1\right)\right)\right]}{\left[P_G \cdot \left(1 - F_{GOOD}\left(S^F\right)\right)\right]},$$

$$P_{3GOOD} = \frac{\left[P_G \cdot \left(F_{GOOD}\left(S_1\right) - F_{GOOD}\left(S^F\right)\right)\right]}{\left[P_G \cdot \left(1 - F_{GOOD}\left(S^F\right)\right)\right]}.$$

For simplicity define the total arrival rate as

$$\lambda_{123} = \lambda_1 + \lambda_2 + \lambda_3.$$

Therefore the expected waiting time for good items in the in-box is

$$W_{GOODINBOX} = P_{1GOOD} \cdot \mathrm{W}_q^{(1)} + P_{2GOOD} \cdot \mathrm{W}_q^{(2)} + P_{3GOOD} \cdot \mathrm{W}_q^{(3)}.$$

And the expected total wait in the system for good items that reach the decision maker is

$$W_{\text{DM GOOD}} = \left(\frac{1}{\mu^F - \lambda_{COLLECT}}\right) + W_{GOODINBOX} + \left(\frac{1}{\mu_{DM}}\right).$$

C. $W_{\text{TYPE II ERROR}}$

The expected waiting time for bad items that reach the decision maker is calculated in the same manner. Define the probability weights as

$$P_{2BAD} = \frac{\left[P_B \cdot \left(1 - F_{BAD}(S_1)\right)\right]}{\left[P_B \cdot \left(1 - F_{BAD}(S^F)\right)\right]}$$

and

$$P_{3BAD} = \frac{\left[P_B \cdot \left(F_{BAD}(S_1) - F_{BAD}(S^F)\right)\right]}{\left[P_B \cdot \left(1 - F_{BAD}(S^F)\right)\right]}.$$

Therefore the expected waiting time for bad items in the in-box is

$$W_{BADINBOX} = P_{2BAD} \cdot \mathrm{W}_q^{(2)} + P_{3BAD} \cdot \mathrm{W}_q^{(3)}.$$

And the expected total wait in the system for bad items that reach the decision maker is

$$W_{\text{TYPE II ERROR}} = \left(\frac{1}{\mu^F - \lambda_{COLLECT}}\right) + W_{BADINBOX} + \left(\frac{1}{\mu_{DM}}\right).$$

D. $W_{\text{TRASH BAD}}$ and $W_{\text{TYPE I ERROR}}$

Good and bad items that are trashed only wait in the first selector. That wait is

$$\frac{1}{\mu_{DM} - \lambda_{DM}}.$$

LIST OF REFERENCES

1. Ross, Sheldon M., *Introduction to Probability Models*, Fifth Edition, Academic Press, San Diego, CA, 1993

2. Woosley, Roy B., "Application of Expert System Technology to an Information Fusion System", *Second International Conference on Command, Control, Communications and Management Information Systems*, 1-3 April 1987

3. Devore, Jay L., *Probability and Statistics for Engineering and the Sciences*, Forth Edition, Duxbury Press, Belmont, CA, 1995

4. Gross, D. and Harris, C. M., *Fundamentals of Queueing Theory*, Second Edition, John Wiley & Sons, New York, NY, 1985

INITIAL DISTRIBUTION LIST

No. Copies

1. Defense Technical Information Center 2
   8725 John J. Kingman Rd., Ste 0944
   Ft. Belvoir, VA 22060-6218

2. Dudley Knox Library 2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, CA 93943-5101

3. Chairman, Code OR 1
   Department of Operations Research
   Naval Postgraduate School
   Monterey, CA 93943-5101

4. Professor Kneale T. Marshall, Code OR/Mt 4
   Department of Operations Research
   Naval Postgraduate School
   Monterey, CA 93943-5101

5. Professor Peter Purdue, Code OR/Pd 2
   Department of Operations Research
   Naval Postgraduate School
   Monterey, CA 93943-5101

6. Center for Operations Research, Code R55 2
   9800 Savage Road
   Fort George G. Meade
   MD 20755-6000

7. LCDR John A. Sears, III 2
   224 South Aberdeen Ave.
   Wayne, PA 19087